National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

NL-339 (r.86/06)

Canada

THE UNIVERSITY OF ALBERTA

STABLE ADAPTIVE CONTROL IN THE PRESENCE OF UNMODELED DYNAMICS

by

WILLIAM R. CLUETT

©

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMICAL ENGINEERING

EDMONTON, ALBERTA

FALL 1986

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-32606-9

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR WILLIAM R. CLUETT

TITLE OF THESIS STABLE ADAPTIVE CONTROL IN THE PRESENCE OF UNMODELED DYNAMICS

DEGREE FOR WHICH THESIS WAS PRESENTED DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED FALL 1986

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) ....

PERMANENT ADDRESS:

486 BROADWAY AVE.

TORONTO, ONTARIO

M4G 2R6

DATED ....Sept. 26......1986

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled STABLE ADAPTIVE CONTROL IN THE PRESENCE OF UNMODELED DYNAMICS submitted by WILLIAM R. CLUETT in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

....

....

Co-supervisors

V. G. Gourishankar

....

....

....

B. Erik Ydstie

....

External Examiner

Date........Sept. 26, 1986........

To Grampy Zinck and Stephen Court

Abstract

This thesis examines the stability of adaptive control systems in the presence of model-plant mismatch. Most stability results for adaptive control systems are based on the assumption that the order of the model used in the control structure is greater than or equal to that of the process. However most real processes are high order and hence, in general, a lower order control structure is used in practice. It has been established in the literature that straightforward application of these 'stable' algorithms may lead to stability problems when unmodeled dynamics, due to this model-process order mismatch, are present.

This work makes three main contributions. Firstly, a design approach for discrete adaptive control systems is presented which provides a quantitative measure of the effect of design alternatives such as (i) adaptive gain, (ii) model order, and (iii) sampling rate, on stability in the presence of unmodeled plant dynamics. The proposed approach, based on conic sector stability theory, is illustrated using a benchmark example. The results demonstrate that the sector conditions permit design tradeoffs to be made such that stability can be maintained despite the presence of model-plant mismatch.

Secondly, a new global stability theorem for adaptive predictive control systems (APCS) is presented which states the requirements for stability in terms of one concise condition. This condition is not problem specific and

depends only on the convergence properties of the *a posteriori* process output estimation error and the estimated controller parameters. It is demonstrated that this new theorem is applicable to a broad class of linear, time-varying, stochastic processes.

Thirdly, the stability of an adaptive predictive control system in the presence of model-process order mismatch is established using this new theorem. The proposed design approach uses a *normalized* parameter estimation scheme, which permits a formal proof that the modeling errors can be treated as a *bounded* disturbance, and a parameter adaptation stopping criterion to guarantee global stability. Simulation results, based on a benchmark example, show that the proposed algorithm resulted in stable performance under conditions where the equivalent algorithm without the normalized estimation scheme and the on/off criterion was unstable. These results are also extended to a generalized minimum variance type of adaptive controller. In this formulation, the plant representation has been augmented to incorporate P, Q and R weighting polynomials for the system output, input and setpoint, respectively into the predictive control law. This algorithm also includes a normalized estimation scheme based on a least squares estimator.

Acknowledgement

I would like to acknowledge the assistance and guidance of my supervisors Dr. S.L. Shah and Dr. D.G. Fisher throughout the course of this thesis. I would also like to thank Dr. J.M. Martin-Sanchez and Dr. J.W. Macki for many valuable discussions and suggestions. I would also like to extend a special thanks to Dr. Shah for encouraging me to tackle the robust adaptive control problem.

I also want to thank my wife, Janet, for her constant love, encouragement and support during this chapter of our lives.

Financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) in the form of a 1967 Science Scholarship, and from the University of Alberta in the form of a Dissertation Fellowship is also gratefully acknowledged.

Table of Contents

Chapter Page

1. Introduction 1

1.1 Scope 1

1.2 Objectives 2

1.3 Outline 4

1.4 References 5

2. Robustness of Continuous-Time Adaptive Controllers ...7

2.1 Introduction 7

2.2 Problem Formulation 7

2.2.1 Error Systems 7

2.2.2 Adaptive Error System 10

2.3 A MRAC Example 13

2.3.1 Choice of the Tuned Controller 16

2.4 Global Stability 22

2.4.1 Continuous-time Systems 22

2.4.2 Discrete-time Systems 25

2.5 Local Stability 28

2.6 Conclusions 33

2.7 References 34

3. Robustness of Discrete-Time Adaptive Controllers 35

3.1 Introduction 35

3.2 Problem Formulation 35

3.2.1 Plant and Controller Definition 35

3.2.2 Error Model 37

3.3 Two Parameter Adaptation Algorithms 40

3.4 Stability Analysis 42

3.4.1 Normalized System 42

3.4.2 Extension to the Original Error System 42

3.5 Conclusions 48

3.6 References 50

4. Robust Design of Adaptive Control Systems 51

4.1 Introduction 51

4.2 Linearization Analysis 53

4.2.1 Rohrs' Error System 53

4.2.2 Rohrs' Design Guidelines 60

4.3 A Comment on Rohrs' Results 65

4.4 A Design Approach for Discrete Adaptive Control Systems 84

4.4.1 A Robust Stability Result 86

4.4.2 Proposed Design Approach 90

4.4.3 Adaptive Gain Effects 94

4.4.4 Model Order Effects 95

4.4.5 Sampling Time Effects 98

4.5 Conclusions 101

4.6 References 102

5. Stability Analysis of Discrete-Time Adaptive Control Systems based on Input-Output Theory 104

5.1 Introduction 104

5.2 Nonzero Initial Conditions 105

5.3 The Vanishing Radius Problem 110

5.4 Choice of the Normalization Factor 113

5.5 Extension to a L_∞ Result 128

5.6 Addition of P and Q Weighting to the Control Law 139

5.7 Conclusions 142

5.8 References 144

6. A New Global Stability Result for the Model-Plant Mismatch Problem 145

6.1 Introduction 145

6.2 General Conditions for Stability 150

6.2.1 Process Description 150

6.2.2 APCS Description 151

6.2.3 Basic Assumptions 152

6.2.4 Global Stability Condition 153

6.2.5 A General Stability Result 155

6.3 Verification of the Stability Condition in the Presence of Unmodeled Dynamics 157

6.3.1 A Normalized Estimation System 158

6.3.2 Convergence Properties of the Normalized System 162

6.3.3 Properties of the Unnormalized a posteriori Estimation Error 166

6.4 Global Stability Result 167

6.5 Simulation Results 168

6.6 Discussion of Results 175

6.7 A Robust Adaptive Controller 178

6.7.1 A Normalized Least Squares Parameter Estimator 180

6.7.2 Convergence Properties of the Normalized System 181

6.7.3 Global Stability Result 183

6.7.4 A Self-Tuning Control Law 185

6.8 Conclusions 187

6.9 References 189

7. Conclusions and Recommendations 192

7.1 Conclusions 192

7.2 Recommendations 196

7.3 References 197

Appendix A: Mathematical Preliminaries 198

List of Figures

Figure Page

2.1 Error system block diagram 14

2.2 Tuned controller structure schematic 17

2.3 Fixed parameter control system block diagram 20

2.4 Martin-Sanchez's feedback system schematic 27

2.5 Feedback system schematic for local stability analysis 29

3.1 Error model schematic 39

3.2 Normalized error model schematic 49

4.1 Rohrs' error system schematic 58

4.2 Rohrs' linearized error system schematic 59

4.3 Discrete plant zero locations versus sample time 68

4.4 Simulation with no unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$ 69

4.5 Simulation with unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$ 70

4.6 Simulation with no unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$ 71

4.7 Simulation with unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$ 72

4.8 Simulation with unmodeled dynamics: $r(t)=1.0+4.5\sin(13.5t)$ and $\gamma=0.1$ 74

4.9 Simulation with no unmodeled dynamics: $r(t)=1.0+4.5\sin(13.5t)$ and $\gamma=0.1$ 75

4.10 Simulation with no unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$ 77

4.11 Simulation with no unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$ 78

4.12 Simulation with no unmodeled dynamics: $r(t)=1.0+4.5\sin(6.75t)$ and $\gamma=0.1$ 79

Figure Page

4.13 Simulation with unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$ 80

4.14 Simulation with unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$ 81

4.15 Simulation with unmodeled dynamics: $r(t)=1.0+4.5\sin(6.75t)$ and $\gamma=0.1$ 82

4.16 Ortega *et al*'s error model schematic 89

4.17 Convergence trajectories for first order model parameters 93

4.18 Nyquist locus of $H_2(\mu^{1/2}q)$ for a first order model ($T_s=.32s$, $\mu=0.9$) 96

4.19 Nyquist locus of $H_2(\mu^{1/2}q)$ ($T_s=0.24s$, $\mu=0.9$) for: (a) a first order model; and (b) a second order model 97

4.20 Nyquist loci of H_2 for first order models 99

6.1 Simulation of A1 with $T_s=0.3s$ 170

6.2 Simulation of A1 with $T_s=0.25s$ 171

6.3 Simulation of A2 with $T_s=0.3s$ 173

6.4 Simulation of A2 with $T_s=0.25s$ 174

Nomenclature

Technical Abbreviations

APCS	Adaptive Predictive Control System
ARMA	Auto Regressive Moving Average
CG	Constant Gain
KTL	Key Technical Lemma
LS	Least Squares
MPM	Model-Plant Mismatch
MRAC	Model Reference Adaptive Control
NCS	Nominally Controlled System
PAA	Parameter Adaptation Algorithm
PE	Persistent Excitation
SISO	Single Input Single Output
SPR	Strictly Positive Real
STC	Self-Tuning Control

Alphabetic

$A(q^{-1})$	polynomial containing process output parameters
$B(q^{-1})$	polynomial containing process input parameters
d	process time delay
e	error signal
n	normalization factor
$P(q^{-1})$	weighting polynomial for process output
q^{-1}	backward shift operator (i.e. $q^{-1}[u(k)]=u(k-1)$)
$Q(q^{-1})$	weighting polynomial for process input
r	reference input signal
$R(q^{-1})$	weighting polynomial for desired output
u	control input signal to process

y	process output signal
z	augmented process output signal

Greek

ζ	regressor vector (continuous-time)
θ	parameter vector
ξ	disturbance input signal to process
ρ	normalization factor
ϕ	regressor vector (discrete-time)

Superscripts

n	normalized signal
t	transpose
$*$	tuned value
$\hat{}$	estimated value
$\sim$	deviation value

Subscripts

d	desired value
m	reference model value

1. Introduction

Globally stable discrete-time adaptive controllers have been available in the literature since 1980 (e.g. Goodwin *et al*, 1980). Most stability results for adaptive control systems are based on the assumption that the order of the model used in the control law is greater than or equal to that of the process. However, most real processes are high order and hence, in general, a lower order control structure is used in practice.

Rohrs (1982) demonstrated that straightforward application of these 'stable' algorithms may lead to stability problems when unmodeled dynamics, due to this model-process order mismatch, are present. Rohrs' (1982) work has sparked a new wave of research into the performance and behaviour of adaptive control systems when the model order assumption is violated.

1.1 Scope

Several other assumptions are commonly used to prove global stability of adaptive control systems. For instance, the assumptions made by Goodwin *et al* (1980) about the plant, in addition to the model order assumption described above, are:

(i) the time delay is known,

(ii) the plant is linear and time-invariant,

(iii) the plant has all of its zeros inside the unit circle,

(iv) the plant is disturbance-free.

This set of assumptions is restrictive and certainly some subset of these assumptions are violated by real processes. Therefore a complete robustness analysis of adaptive control schemes should examine the question of stability under more realistic conditions (e.g. non-linear, time-varying plants). The scope of this thesis is restricted to the stability analysis of adaptive control systems in the presence of model-plant mismatch. Since the majority of adaptive algorithms are implemented on a digital computer, the work in this thesis is based primarily on discrete-time algorithms.

1.2 Objectives

If adaptive controllers are to find their way into practical and commercial applications, then the robustness issues must receive further attention. It is well known that the assumptions made regarding the plant are violated to some degree in practice. However, several successful attempts to apply adaptive controllers have been reported. Many of these applications have included engineering modifications to the basic algorithms in order to improve robustness and performance of the controller. Despite these promising results, there exists little rigorous theory to guide the engineer from the simple adaptive strategies through to an effective and robust scheme. The overall objective of this thesis is to contribute to the understanding and development of a theoretical basis for the design of robust adaptive controllers.

One of the first assumptions to be scrutinized in the literature was the issue of stability in the presence of noise and disturbances. The algorithms presented by Egardt (1979), Martin-Sanchez (1982) and Samson (1983) consider the existence of such disturbances. All three of these algorithms include a mechanism for turning parameter adaptation on and off which is used in the proof of global stability. The analysis presented by Martin-Sanchez (1982) defines a perturbation signal which includes the effect of noise and disturbances on the process output. Martin-Sanchez *et al* (1984) suggested that unmodeled dynamics may be included in the perturbation signal. Prompted by this proposal, the first main objective of this thesis was to examine the role of an on/off parameter adaptation mechanism for adaptive control systems in the presence of unmodeled plant dynamics.

Praly (1984) and Ortega *et al* (1985) in their robustness analyses modified the adaptive control algorithm by normalizing the signals entering the parameter adaptation scheme. The second objective of this thesis was to investigate the role of normalization and its importance for robustness of adaptive controllers in the presence of modeling errors.

Rohrs *et al* (1984) presented some guidelines for the design of discrete-time adaptive controllers which improve stability in the presence of unmodeled dynamics. These guidelines were extracted from a linearization analysis of the inherently nonlinear adaptive control problem. The third objective of this thesis was to analyze and attempt to justify these guidelines from a more rigorous theoretical perspective.

Although not completely documented as part of this thesis, this project involved a significant amount of background study in stability theory, stability of adaptive control systems, and current robustness literature found mostly in recent conference proceedings.

1.3 Outline

This thesis consists of seven chapters. Chapters two and three provide a clarification of work published by other authors (Kosut and Johnson, 1984; Ortega *et al*, 1985) in the area of robust adaptive control. This work was selected because the authors have cast the problem into a well formulated, adaptive error feedback system and have applied standard input-output stability results. Chapter four describes some of Rohrs (1982) work in adaptive control and presents some significant extensions. Chapter five examines further the input-output stability approach to analyzing robustness in the presence of unmodeled dynamics. Chapter six presents a new stability theorem for adaptive control systems and uses this theorem to prove global stability of an adaptive predictive control system in the presence of model-process mismatch. The thesis follows the paper format and therefore each chapter may be read independently. The overall conclusions and recommendations for future work are given in Chapter seven.

1.4 References

Egardt, B. (1979). *Stability of Adaptive Controllers*, Springer-Verlag, Berlin.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Kosut, R.L. and C.R. Johnson, Jr. (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Martin-Sanchez, J.M. (1982). A globally stable APCS in the presence of bounded noises and disturbances. *Proc. 21st IEEE CDC*, 761-769; also in *IEEE Trans. Aut. Contr.*, AC-29, 461-464, 1984.

Martin-Sanchez, J.M., S.L. Shah and D.G. Fisher (1984). A stable adaptive predictive control system. *Int. J. Contr.*, 39, 215-234.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Praly, L. (1984). Robust model reference adaptive controllers, part 1: stability analysis. *Proc. 23rd IEEE CDC*, 1009-1014.

Rohrs, C. (1982). Adaptive control in the presence of unmodeled dynamics. Dept. of Elec. Eng. Comput. Sci., MIT, LIDS Report No. TH-1254.

Rohrs, C.E., M. Athans, L. Valavani and G. Stein (1984). Some design guidelines for discrete-time adaptive controllers. *Automatica*, 20, 653-660.

Samson, C. (1983). Stability analysis of adaptively controlled systems subject to bounded disturbances. *Automatica*, 19, 81-86.

2. Robustness of Continuous-Time Adaptive Controllers

2.1 Introduction

This chapter is a tutorial/discussion based primarily on the work of R. L. Kosut and co-workers. Their recent research in the area of robust adaptive control can be found in three major papers (Kosut and Johnson, 1984; Kosut and Friedlander, 1985; Kosut and Anderson, 1986). Kosut and Johnson (1984) present a concise summary of the global and local stability results in Kosut and Friedlander (1985) and Kosut and Anderson (1986), respectively. The major tool used throughout these papers for studying the robustness of continuous-time adaptive control systems is input-output stability theory. The framework for the work is an adaptive error system based on the concept of a tuned system. An excellent reference for the stability theory is Desoér and Vidyasagar (1975).

2.2 Problem Formulation

2.2.1 Error Systems

Kosut and Johnson (1984) begin with the formulation of an error system. First a very general expression is written for the error output

$$e=H_{ew}(\theta)\cdot w \qquad (2.1)$$

where H_{ew} is an operator which represents the closed loop relationship between external inputs w and the error e. $H_{ew}(\cdot)$ is a function of the controller parameters θ. If the controller was adaptive, it would adjust the parameters on-line in order to reduce the error. In a nonadaptive setting, the designer of the control system would pick a set of parameters θ which would satisfy a certain objective. As Kosut and Johnson (1984) point out, this set may not be unique. In other words, there could be several θ which satisfy the designer's objectives.

Kosut and Johnson (1984) present three parameter sets which correspond to three distinct objectives. First is the *matched* parameter set $\hat{S}$.

$$\hat{S}=\{\hat{\theta} : H_{ew}(\hat{\theta})=0\} \tag{2.2}$$

This would imply that the error output for all external inputs is identically zero for all time. Next is the *robust* parameter set S^0.

$$S^0=\{\theta^0 : ||H_{ew}(\theta^0)\cdot w||/||w|| \le p^0, \forall\ w\epsilon W\}$$

Note that the bound p^0 applies to all sets of inputs. The third parameter set is the *tuned* set $S_w{}^*$.

$$S_w{}^*=\{\theta_w{}^* : ||H_{ew}(\theta_w{}^*)\cdot w||/||w|| \le p^*\} \tag{2.3}$$

Note that this set applies to a particular $w\epsilon W$. For the tuned system to be meaningful with respect to the robust set

$$p^*<p^0 \tag{2.4}$$

i.e. the acceptable bound for the robust set is less restrictive than that for the tuned set. This is because the robust set must accomodate all $w\epsilon W$ whereas the tuned set applies only to a specific w.

Kosut and Johnson (1984) also define a more general tuned set of controller parameters S^*

$$S^* = \bigcup_{w \in W} S_w^* \tag{2.5}$$

where each element of S^* is a tuned set for a particular w. However there is no restriction that members of a subset S_w^* must provide any sort of satisfactory control for a different w. (Membership in S^* is denoted by $\theta^* \epsilon S^*$ where the dependence on w is implied.)

The matched case has been the focus of stability research in adaptive control. For instance, many algorithms in the literature assume deterministic plants where an upper bound on the order of the plant is known. Therefore there exists a fixed set of parameters (ie. $\hat{\theta}$) which provides a zero error. These assumptions are surely violated in the case of a real plant because of the presence of unmeasured disturbances and unmodeled dynamics. In this case, it would not be possible to find a set of parameters to solve $H_{ew}(\hat{\theta})=0$. The other two sets of parameters, S^0 and S_w^*, are more realistic in their requirements. Neither set requires that $e=0$ for all time but only a bound on the norm of the error is necessary.

From equation (2.1), the tuned error and robust error are defined as

$$e^* = H_{ew}(\theta^*) \cdot w \tag{2.6}$$

$$e^0 = H_{ew}(\theta^0) \cdot w \tag{2.7}$$

From equation (2.4), it follows that for a particular $w \epsilon W$,

$$||e^*|| < ||e^0||$$

Adaptive control can be justified if, for a large subset of $w \epsilon W$, various tuned parameter sets exist such that,

$$||e^*|| << ||e^0|| \tag{2.8}$$

Otherwise, the fixed parameter robust controller would be quite adequate.

The above discussion is quite a simple and elegant justification for adaptive control. It is worth noting that there is no unique set of parameters to which the controller might converge. Instead there is a union of parameter sets available which satisfy the *tuned* criterion. Kosut and Johnson (1984) use this tuned set concept to develop a generic adaptive error system.

2.2.2 Adaptive Error System

Kosut and Johnson (1984) consider an adaptive version of (2.1), as described by the following equations.

$$\begin{bmatrix} e \\ \\ \zeta \end{bmatrix} = \begin{bmatrix} H_{ew}(\hat{\theta}) \\ \\ H_{\zeta w}(\hat{\theta}) \end{bmatrix} \cdot w = H(\hat{\theta}) \cdot w \tag{2.9}$$

$$\hat{\theta} = \Omega \{\hat{\theta}(0), e, \zeta\} \tag{2.10}$$

where $\hat{\theta}$ is the estimated controller parameter vector which is updated by the parameter adaptation algorithm (PAA) Ω and $\hat{\theta}(0)$ is the initial estimate of θ. The PAA is a function of the error signal and the regressor vector ζ which contains information from signals within the adaptive system.

The objective of Kosut and Johnson's (1984) analysis is to determine the conditions under which the system in (2.9) and (2.10) is stably attracted to the set of tuned systems. In order to formulate the error system, the structure of the control law must be selected. Kosut and Johnson (1984) select the well-known bilinear expression

$$u(t)=-\zeta^{t}(t)\hat{\theta}(t) \tag{2.11}$$

where u is the control input to the plant. An assumption is made at this point with regards to the way in which u and w are mapped through H into e and ζ.

Assumption: The map $(w,u)\rightarrow(e,\zeta)$ is linear, time-invariant, i.e.

$$\begin{bmatrix} e(t) \\ \zeta(t) \end{bmatrix} = \begin{bmatrix} G_{ew}(s) & G_{eu}(s) \\ G_{\zeta w}(s) & G_{\zeta u}(s) \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix} = G(s) \begin{bmatrix} w(t) \\ u(t) \end{bmatrix} \tag{2.12}$$

where G(s) is the open-loop interconnection matrix whose elements are proper, rational functions and s denotes the Laplace transform variable.

Kosut and Johnson (1984) further define the parameter error, $\tilde{\theta}$, and the adaptive control error, v, as

$$\tilde{\theta}(t)=\hat{\theta}(t)-\theta^{*} \qquad \theta^{*}\epsilon S^{*} \tag{2.13}$$

and

$$v(t)=\zeta^{t}(t)\tilde{\theta}(t) \tag{2.14}$$

Substituting (2.11) into (2.12) and using the definition of v, gives

$$\begin{bmatrix} e(t) \\ \zeta(t) \end{bmatrix} = -G(s) \begin{bmatrix} 0 \\ \zeta^t(t)\theta^* \end{bmatrix} + G(s) \begin{bmatrix} w(t) \\ -v(t) \end{bmatrix} \qquad (2.15)$$

Equation (2.15) may be rewritten in a more compact form

$$\begin{bmatrix} e(t) \\ \\ \zeta(t) \end{bmatrix} = \begin{bmatrix} H_{ew}{}^* & H_{ev}{}^* \\ \\ H_{\zeta w}{}^* & H_{\zeta v}{}^* \end{bmatrix} \begin{bmatrix} w(t) \\ \\ -v(t) \end{bmatrix} = H^* \begin{bmatrix} w(t) \\ \\ -v(t) \end{bmatrix} \qquad (2.16)$$

where

$$H^* = f(G,\theta^*) \qquad (2.17)$$

Kosut and Johnson (1984) give the exact relationships implied in (2.17) but do not explicitly use these expressions in their later analysis of a model reference adaptive control (MRAC) system. More important is the general error system which follows from (2.16). Kosut and Johnson (1984) define the tuned error, e^*, and tuned regressor, ζ^*, as

$$e^*(t)=H_{ew}{}^* \cdot w(t) \qquad (2.18)$$

$$\zeta^*(t)=H_{\zeta w}{}^* \cdot w(t) \qquad (2.19)$$

Then the error system may be expressed as

$$e(t)=e^*(t)-H_{ev}^*\cdot v(t) \tag{2.20}$$

$$\zeta(t)=\zeta^*(t)-H_{\zeta v}^*\cdot v(t) \tag{2.21}$$

$$v(t)=\zeta^t(t)\tilde{\theta}(t) \tag{2.22}$$

$$\tilde{\theta}(t)=\Omega[\tilde{\theta}(0),e,\zeta] \tag{2.23}$$

This error system is depicted in Figure 2.1 and consists of a nonlinear block and a linear time-invariant block in a feedback arrangement. This system is disturbed by e^* which enters additively and ζ^* and $\tilde{\theta}(0)$ which disturb the nonlinear block.

Kosut and Johnson (1984) do an analysis of a particular continuous-time controller. This example will be reviewed because it gives some very helpful insight into the error system approach to studying the stability problem.

2.3 A MRAC Example

Kosut and Johnson (1984) select for further robustness analysis the well-known model reference approach to adaptive control, i.e. the control objective is to have the plant match the behaviour of a user supplied reference model. The following equations describe the plant and model.

$$y(t)=P(s)u(t)+\xi(t) \quad \text{(plant)} \tag{2.24}$$

$$y_m(t)=\bar{H}(s)r(t) \quad \text{(reference model)} \tag{2.25}$$

$$e(t)=y(t)-y_m(t) \quad \text{(tracking error)} \tag{2.26}$$

where ξ is a disturbance term and r is the reference signal.

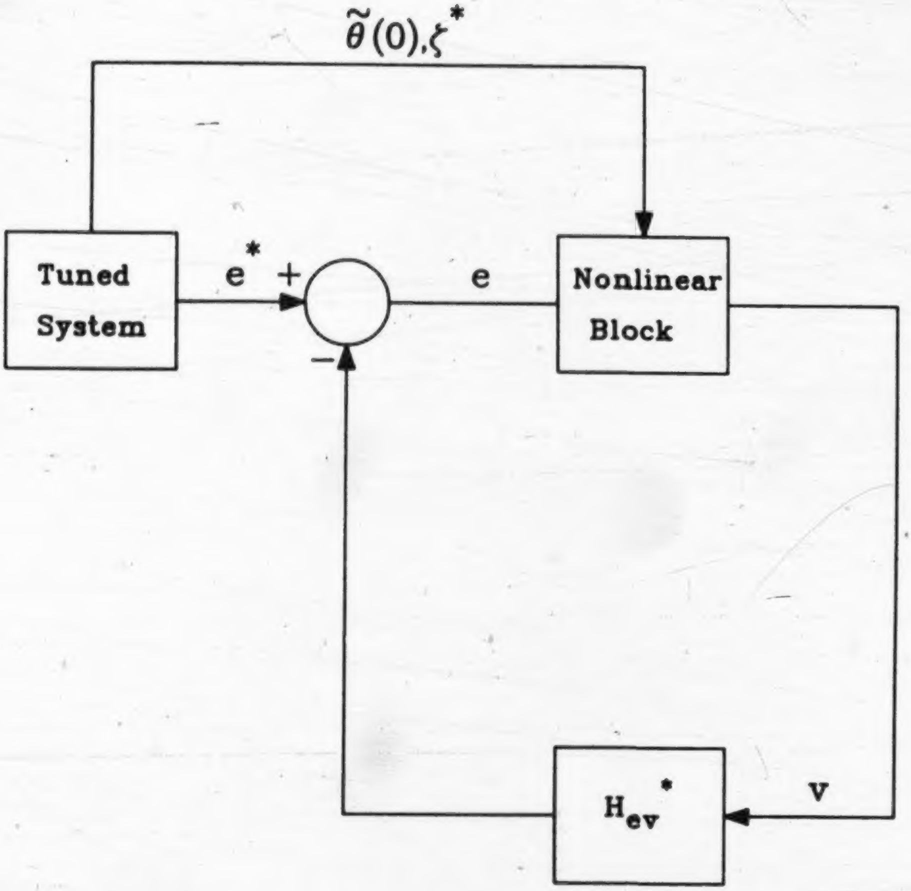
$\tilde{\theta}(0),\zeta^*$
Tuned System
e^*
+
−
e
Nonlinear Block
v
H_{ev}^*

Figure 2.1 Error system block diagram

Thus the control objective is to make the tracking error *as small as possible*. Kosut and Johnson (1984) use the notation $\{C,\Omega\}$ to represent the adaptive controller where Ω is the PAA and C is the control law. Using the bilinear form, C can be written as

$$u(t)=-\zeta^t(t)\hat{\theta}(t) \tag{2.27}$$

where the regressor contains filtered values of u, y and r.

$$\zeta^t(t)=[\zeta_u{}^t(t),\zeta_y{}^t(t),\zeta_r{}^t(t)]$$

$$=[F(s)u(t),F(s)y(t),-F(s)r(t)] \tag{2.28}$$

with

$$F(s)=[1/L(s),\cdots,s^{k-1}/L(s)] \tag{2.29}$$

$$L(s)=s^k+a_1s^{k-1}+\cdots+a_k \tag{2.30}$$

For the system of (2.24-2.26), the disturbance w in (2.16) is given by

$$w^t(t)=[r(t)\quad \xi(t)] \tag{2.31}$$

The plant and model expression of (2.24-2.26) with the control law C of (2.27-2.30) are used by Kosut and Johnson (1984) to develop the tuned signals, e^* and ζ^*, and the tuned interconnections, $H_{ev}{}^*$ and $H_{\zeta v}{}^*$ in (2.20) and (2.21). These results are omitted here because they are an intermediate step towards the final tuned system used for stability analysis. Kosut and Johnson (1984) select the following plant representation

$$P(s)=[1+\Delta(s)]P^*(s) \tag{2.32}$$

$$P^*(s)=b_0B^*(s)/A^*(s) \qquad (2.33)$$

where A^* and B^* are polynomials of order n and m respectively with $n>m$. P^* is referred to as the *tuned parametric model* which would be a good representation of P at perhaps low frequencies. The transfer function for Δ represents the unmodeled dynamics and would account for the dynamics not contained in P^* say at high frequencies. According to Doyle and Stein (1981), Δ is referred to as *model uncertainty* and (2.32) is in a *multiplicative form*. The structure of Δ is not known but it is assumed to be stable and bounded as follows

$$|\Delta(j\omega)| \le \delta(\omega) \qquad \forall\ \omega\epsilon R \qquad (2.34)$$

The reference model is selected to be

$$\bar{H}(s)=\bar{B}(s)/\bar{A}(s) \qquad (2.35)$$

where $\bar{A}$ and $\bar{B}$ are polynomials of order n and m respectively and $\bar{H}$ is stable.

2.3.1 Choice of the Tuned Controller

To proceed any further, some decision must be made with regards to the choice of the tuned controller. This step has a large impact on the analysis because, as will be seen below, the tuned signals and interconnections become functions of the tuned controller parameters. Kosut and Johnson (1984) select the control law of Egardt (1979) for the structure of their tuned controller. Figure 2.2 is taken from Egardt (1979). From this figure it can be readily seen that

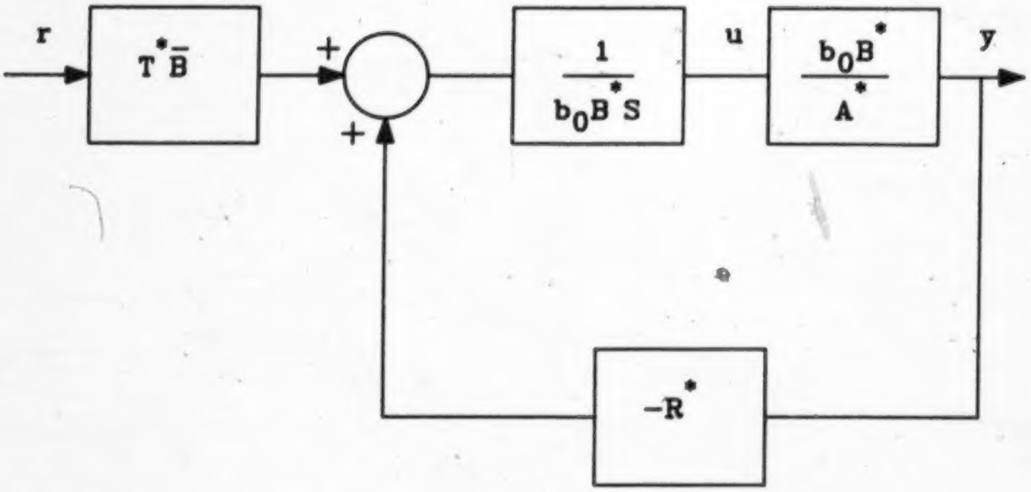
r
$T^*\bar{B}$
+
+
$\frac{1}{b_0B^*S}$
u
$\frac{b_0B^*}{A^*}$
y
$-R^*$

Figure 2.2 Tuned controller structure schematic

$$u(t)=(b_0B^*S^*)^{-1}[T^*\bar{B}r(t)-R^*y(t)] \tag{2.36}$$

where

$$T^*\bar{A}=A^*S^* + R^* \tag{2.37}$$

The choice of control law (2.36) gives an exact match between the process output and the model output when the process is completely described by P^* (2.33) and there are no disturbances (i.e. $\xi=0$). Therefore this parameter set is a member of the matched set $\bar{S}$ for the tuned model.

By inserting the adaptive control error v in (2.22) into the control law of (2.27-2.30), (2.27) becomes

$$u(t)=C_{ur}{}^*(s)r(t)-C_{uy}{}^*(s)y(t)-C_{uv}{}^*(s)v(t) \tag{2.38}$$

$C^*=[C_{ur}{}^*,C_{uy}{}^*,C_{uv}{}^*]$ is referred to as the tuned controller. Comparing (2.38) and (2.36) gives the following expressions

$$C_{uy}{}^* = R^*/(b_0B^*S^*) \tag{2.39}$$

$$C_{ur}{}^* = \bar{B}T^*/(b_0B^*S^*) \tag{2.40}$$

Equations (2.39) and (2.40) ensure that, when $v(t)=0$ and $\xi(t)=0$, C^* will provide exact model following (i.e. $e(t)=0$) with the P^* process.

Kosut and Johnson (1984) state that the tuned design must be robust. The properties of this particular choice for C^* are stated above. The selection of C^* impacts the properties of the tuned signals and tuned interconnections as will be seen below. In Kosut and Johnson (1984) it is shown that

$$C_{uv}{}^* = L/(B^*S^*) \tag{2.41}$$

It is also convenient to define

$$G^* = R^*/(T^*\bar{A}) \tag{2.42}$$

Using equations (2.39-2.41), the tuned signals of (2.20) and (2.21) are given by

$$e^* = (1+\Delta G^*)^{-1}[\Delta(1-G^*)\bar{H}r + A^*S^*/(\bar{A}T^*)\cdot\xi] \tag{2.43}$$

$$\zeta^* = \begin{bmatrix} (1+\Delta G^*)^{-1}[\frac{A^*\bar{B}}{b_0 B^*\bar{A}}F^t r - \frac{A^*R^*}{b_0 B^* T^*\bar{A}}F^t\xi] \\ (1+\Delta G^*)^{-1}[(1+\Delta)\frac{\bar{B}}{\bar{A}}F^t r - \frac{S^*A^*}{T^*\bar{A}}F^t\xi] \\ -F^t r \end{bmatrix} \tag{2.44}$$

and the tuned interconnections are

$$H_{ev}^* = (1+\Delta G^*)^{-1}(1+\Delta)\frac{b_0 L}{T^*\bar{A}} \tag{2.45}$$

$$H_{\zeta v}^* = \begin{bmatrix} (1+\Delta G^*)^{-1}\frac{A^*L}{B^*T^*\bar{A}}F^t \\ (1+\Delta G^*)^{-1}(1+\Delta)\frac{b_0 L}{T^*\bar{A}}F^t \\ 0 \end{bmatrix} \tag{2.46}$$

Before getting involved with the stability analysis of the error system (2.20-2.23) it is worth considering just the tuned system with input w and outputs e^* and ζ^* (see Figure 2.3). This system consists of a fixed parameter controller (i.e. C^*) and the plant P. If there is no model error (Δ=0), the tuned error (2.43) becomes

$$e^* = \frac{A^*S^*}{\bar{A}T^*}\cdot\xi \tag{2.47}$$

It can be seen from (2.47) that e^* is not a function of the

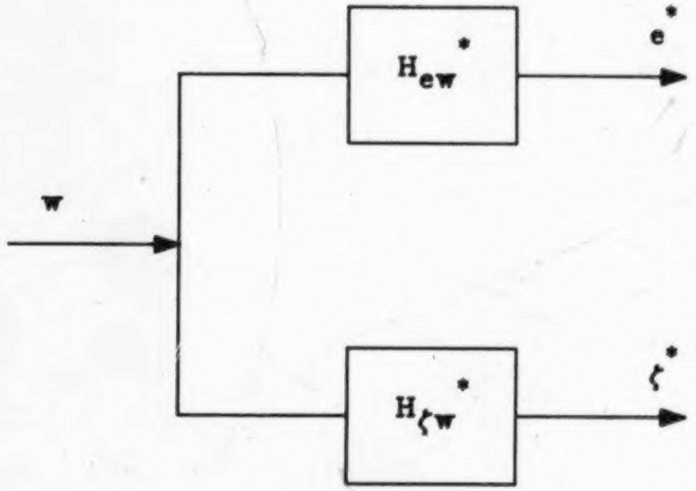
H_{ew}^*
e^*
w
$H_{\zeta w}^*$
ζ^*

Figure 2.3 Fixed parameter control system block diagram

reference signal r in this case and if $\xi=0$, then $e^*=0$. This is in agreement with the remarks made before regarding the properties of the tuned controller. The error output for the case when $\xi\neq0$, is always bounded for bounded disturbances because $\tilde{A}$ and T^* are stable polynomials by assumption. This is also true of the tuned regressor, i.e. ζ^* (2.44) is bounded for bounded r and ξ because B^* is stable by assumption. Therefore the tuned system with no model error is stable. If $\Delta\neq0$, the *actual* tuned system is stable if and only if:

$$(1+\Delta G^*)^{-1} \text{ and } \Delta(1+\Delta G^*)^{-1} \text{ are stable.}$$

It is not necessary that a complete description of Δ is available in order to verify the stability requirements. If Δ is stable, stability of the tuned system will be guaranteed if (e.g. Doyle and Stein, 1981)

$$|\Delta(j\omega)|\cdot|G^*(j\omega)| < 1 \quad \forall\ \omega\epsilon R \tag{2.48}$$

From (2.34), (2.48) may be expressed as

$$|\Delta(j\omega)| < \delta(\omega) = 1/|G^*(j\omega)| \quad \forall\ \omega\epsilon R \tag{2.49}$$

This concludes the summary of Kosut and Johnson's (1984) framework for analyzing the stability of continuous-time adaptive control systems in the presence of unmodeled dynamics and disturbances. The remaining part of this chapter summarizes the global and local stability results in Kosut and Johnson (1984).

2.4 Global Stability

2.4.1 Continuous-time Systems

This subsection discusses global stability conditions for the error system of subsection 2.2.2. In section 2.3 an error system has been summarized for a model reference adaptive control system. Kosut and Johnson (1984) select two adaptive algorithms (i.e. Ω in (2.23)) to study the global stability of the error system. The first is the constant gain algorithm of Narendra *et al* (1980) and the second is the retarded gain algorithm of Kreisselmeier and Narendra (1982). The results from the constant gain algorithm will be discussed in more detail.

Constant Gain Adaptive Algorithm

$$\dot{\tilde{\theta}}(t) = \gamma \zeta(t) e(t) \tag{2.50}$$

where $\gamma > 0$ is the constant adaptive gain. The global stability of the system in (2.20-2.23) (see Figure 2.1) with the expressions in (2.43-2.46) for the continuous model reference adaptive control system is the objective. Kosut and Johnson's (1984) objective in this context of *global* stability is to require as little *a priori* knowledge about the plant and the external inputs as possible to prove stability. The basis for the stability analysis is the passivity theorem of Desoer and Vidyasagar (1975). As shown below, the map $e \to v$ is passive for the algorithm presented in (2.50). Therefore, if H_{ev}^* is strictly positive real (SPR), then certain stability properties follow.

Define the inner product $<x,y>_T$ of elements $x,y \in L_{2e}$ by

$$<x,y>_T = \int_0^T x^t(t)y(t)\, dt$$

Using the algorithm in (2.50), from Kosut and Friedlander (1985)

$$\begin{aligned} <e,v>_T &= <e,\zeta^t\tilde{\theta}>_T \\ &= <\zeta e,\tilde{\theta}>_T \\ &= <\dot{\tilde{\theta}},\tilde{\theta}>_T \qquad \text{(assuming } \gamma=1\text{)} \\ &= 1/2\cdot||\tilde{\theta}(T)||^2-1/2\cdot||\tilde{\theta}(0)||^2 \\ &\geq -1/2\cdot||\tilde{\theta}(0)||^2 \end{aligned}$$

Therefore the map $e \rightarrow v$ is passive. Next $H_{ev}{}^*$ is assumed to be SPR, i.e. $H_{ev}{}^*$ is strictly proper, exponentially stable and there exists a finite constant $\rho>0$ such that

$$\mathrm{Re}[H_{ev}{}^*(j\omega)] \geq \rho|H_{ev}{}^*(j\omega)|^2 \qquad \forall\ \omega \in [0,\infty)$$

Also $H_{\zeta v}{}^*$ is assumed strictly proper and exponentially stable. Kosut and Johnson (1984) then consider two types of inputs to the error system

$$W_0{}^*=\{e^*,\ \zeta^*,\ \tilde{\theta}(0)|e^*,\ \dot{e}^* \in L_2,\ \zeta^* \in L_\infty,\ \tilde{\theta}(0) \in R^P\}$$

$$W_B{}^*=\{e^*,\ \zeta^*,\ \tilde{\theta}(0)|e^* \in L_\infty,\ \zeta^* \in L_\infty,\ \tilde{\theta}(0) \in R^P\}$$

Under the above-mentioned assumptions, algorithm (2.50) results in the following properties

(i) If $(e^*,\ \zeta^*,\ \tilde{\theta}(0)) \in W_0{}^*$ then $\tilde{\theta}$, e, ζ and v are in L_∞.

(ii) If $(e^*,\ \zeta^*,\ \tilde{\theta}(0)) \in W_B{}^*$ and $\zeta \in PE$, then $\tilde{\theta}$, e, ζ and v are in L_∞.

The SPR conditions required for $H_{ev}{}^*$ are difficult to satisfy. A necessary property for $H_{ev}{}^* \in SPR$ is that $H_{ev}{}^*$ have a relative degree of one. However, this requires that the relative degree of the plant is known. Since unmodeled dynamics are considered to be present, this information is not available. Another condition necessary for $H_{ev}^* \epsilon SPR$ is that, if Δ in (2.45) is exponentially stable, $|\Delta(j\omega)|$ must be less than one. This condition is almost certainly violated due to the presence of unmodeled dynamics. Therefore it appears that the SPR assumption is extremely restrictive.

In property (ii), an additional requirement to the W_B^* membership (i.e. bounded reference signals and disturbances) is that ζ satisfy a persistent excitation (PE) requirement. $\zeta(t)$ is PE if there exists positive constants a_1, a_2 and a_3 such that

$$\underline{a_2} I \ge \int_s^{s+a_3} \zeta(t)\zeta(t)^t dt \ge a_1 I \qquad \forall\ s \epsilon R_+$$

This PE condition is not well understood from an implementation point of view because with bounded disturbances it is not known how to guarantee $\zeta \epsilon PE$ since ζ is generated inside the adaptive error system. Kosut and Johnson (1984) also suggest that persistently exciting signals result in a deterioration of any setpoint regulation.

The above summary of the global stability results presents a very strong case against going the global stability route to analyze robustness issues. Kosut and Johnson (1984) use this argument as a lead in to and justification for the establishment of *local* stability results. This subject will be reviewed in section 2.5.

2.4.2 Discrete-time Systems

Before continuing, some mention of applicability of these results to discrete-time systems is necessary. Kosut and Friedlander (1985) state that their global stability theorems carry over virtually intact to discrete systems. This statement will be tested by first determining if the map e→v is passive for some discrete adaptive controller. Choose for example the projection algorithm of Goodwin *et al* (1980). For this control scheme

$$e(k)=y(k)-r(k)$$

$$v(k)=\phi(k-1)^t\tilde{\theta}(k-1)$$

$$\begin{aligned}<e(k),v(k)>_T&=<e(k),\phi(k-1)^t\tilde{\theta}(k-1)>_T\\&=<\phi(k-1)e(k),\tilde{\theta}(k-1)>_T\end{aligned}$$

For the projection algorithm

$$\tilde{\theta}(k)=\tilde{\theta}(k-1)+\frac{a(k)\phi(k-1)e(k)}{1+\phi(k-1)^t\phi(k-1)}$$

Let f be defined as

$$f(k)=a(k)\cdot\{1+\phi(k-1)^t\phi(k-1)\}^{-1}$$

Therefore

$$\begin{aligned}<e(k),v(k)>_T&=<(\tilde{\theta}(k)-\tilde{\theta}(k-1))\cdot f(k)^{-1},\tilde{\theta}(k-1)>_T\\&=\sum_{k=0}^{T}f(k)^{-1}(\tilde{\theta}(k)^t\tilde{\theta}(k-1)-\tilde{\theta}(k-1)^t\tilde{\theta}(k-1))\end{aligned}$$

However, there is no guarantee that $\tilde{\theta}(k)^t\tilde{\theta}(k-1)$ is a non-negative scalar variable. Therefore it can not be concluded that e→v is passive. This represents a breakdown in the carry over to discrete systems.

It was claimed by Martin-Sanchez (1976) that a stability solution had been found for a discrete adaptive control system without unmodeled dynamics and disturbances based on Popov's stability theorem. It can be shown that Popov's theorem as applied by Martin-Sanchez (1976) is very similar to the passivity theorem. Therefore it would appear that this solution overcomes the difficulty of using this stability approach experienced above.

In Martin-Sanchez (1976), it is shown that the nonlinear block in Figure 2.4 is passive. A projection PAA, similar to the one used by Goodwin *et al* (1980), satisfies the passivity condition for the map $s \to -s$ where s is given by

$$\begin{aligned} s(k) &= y(k)-\phi(k-1)^t\hat{\theta}(k) \qquad \text{(assuming delay=1)} \\ &= \phi(k-1)^t\theta-\phi(k-1)^t\hat{\theta}(k) \\ &= -\phi(k-1)^t\tilde{\theta}(k) \end{aligned}$$

This parameter may be referred to as the *a posteriori* estimation error. From the passivity theorem, it may be concluded that the feedback system of Martin-Sanchez (1976) is stable, i.e. $s(k)\to 0$. It may be shown that

$$s(k) = e(k)\cdot\{1+\phi(k-1)^t\phi(k-1)\}^{-1} \tag{2.51}$$

where $e(k)=y(k)-r(k)$. Unless ϕ in (2.51) is *a priori* assumed to be bounded, $s(k)\to 0$ does not necessarily imply that $e(k)\to 0$. Thus the passivity theorem on its own does not prove stability of the tracking error e(k). Additional results were later combined with Martin-Sanchez's (1976) result (i.e. Goodwin *et al*'s (1980) Key Technical Lemma) to prove global stability. Ortega *et al* (1985) have extended these global stability results based on input-output stability theory to discrete systems with unmodeled dynamics and

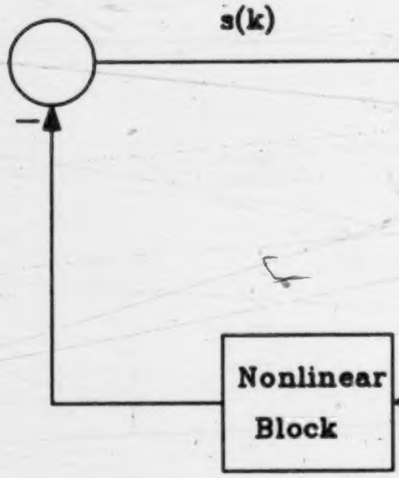
s(k)
−
Nonlinear
Block

Figure 2.4 Martin-Sanchez's feedback system schematic

output disturbances.

2.5 Local Stability

As stated in section 2.4, the conditions required to guarantee global stability are very restrictive and difficult to satisfy. In Kosut and Johnson (1984) and Kosut and Anderson (1986), some *local* stability results are presented. These results put more restrictions on the external input signals than in the global case in order that the plant requirements are less severe. The term *local* is used in the sense of these input restrictions. The development of the local stability analysis is based on the continuous-time error system found in (2.20-2.23) and the constant gain adaptive algorithm of (2.50).

The objective is to transform (2.20-2.23) into a full variational form.

$$\tilde{x} = \tilde{X}_L - \tilde{x}_{NL} \tag{2.52}$$

$$\tilde{x}_{NL} = F \cdot f(\tilde{x}) \tag{2.53}$$

where

$$\tilde{x} = (\tilde{\theta}, \tilde{e}, \tilde{\zeta})$$

$$= (\hat{\theta} - \theta^*,\ e - e^*,\ \zeta - \zeta^*)$$

$$\tilde{x}_L = (\tilde{\theta}_L,\ \tilde{e}_L,\ \tilde{\zeta}_L)$$

$$f(\tilde{x}) = (\tilde{\zeta}^t \tilde{\theta},\ \tilde{\zeta}\tilde{e})$$

This variational form is shown schematically in Figure 2.5. The details for deriving the exact expressions in (2.52) are given by Kosut and Anderson (1986) and will be summarized here. Restating (2.50)

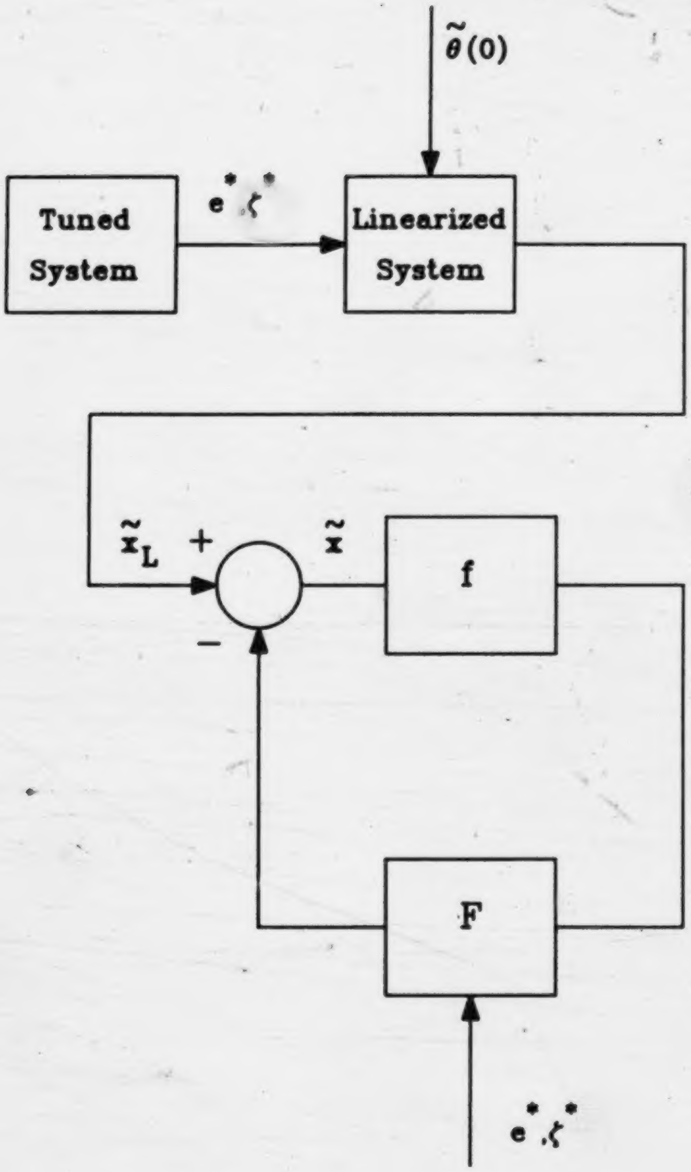
$\tilde{\theta}(0)$
Tuned System
e^*, ζ^*
Linearized System
$\tilde{x}_L$
+
$\tilde{x}$
f
−
F
e^*, ζ^*

Figure 2.5 Feedback system schematic for local stability analysis

$$\dot{\tilde{\theta}}(t) = \gamma\zeta(t)e(t)$$

This may be rewritten as

$$\tilde{\theta} = \tilde{\theta}_0 + L\zeta e \tag{2.54}$$

where $L=(1/s)\gamma$ and s is being used here as the differential operator. Next, the expressions for e and ζ from (2.20) and (2.21) are substituted into (2.54).

$$\tilde{\theta} = \tilde{\theta}_0 + L(\zeta^*-H_{\zeta v}{}^*v)(e^*-H_{ev}{}^*v) \tag{2.55}$$

From (2.22)

$$v = \zeta^t\tilde{\theta} = \tilde{\zeta}^t\tilde{\theta} + \zeta^{*t}\tilde{\theta} \tag{2.56}$$

Substituting (2.56) into (2.55) gives an expression for $\tilde{\theta}$ in terms of $f(\tilde{x})$ and a lumped term which represents $\tilde{\theta}_L$ which is a function only of the tuned signals.

$$\tilde{\theta} = \tilde{\theta}_L - [KN \quad -K][\tilde{\zeta}^t\tilde{\theta} \quad \tilde{\zeta}\tilde{e}]^t \tag{2.57}$$

where

$$\tilde{\theta}_L = [I+LM]^{-1}\tilde{\theta}_0 + K\zeta^*e^*$$

$$N=H_{ev}{}^*\zeta^* + e^*H_{\zeta v}{}^*$$

$$M=N\zeta^{*t}$$

$$K=[I+LM]^{-1}L$$

Expressions for $\tilde{e}$ and $\tilde{\zeta}$ are derived by substituting (2.56) into (2.20) and (2.21) and replacing $\tilde{\theta}$ by (2.57). The results are summarized below.

$$\tilde{x}_L = \begin{bmatrix} [I+LM]^{-1}\tilde{\theta}_0 + K\zeta^*e^* \\ -H_{ev}{}^*\ \zeta^{*t}\tilde{\theta}_L \\ -H_{\zeta v}{}^*\ \zeta^{*t}\tilde{\theta}_L \end{bmatrix} \tag{2.58}$$

$$F = \begin{bmatrix} KN & -K \\ H_{ev}{}^{*}(1-\zeta^{*t}KN) & H_{ev}{}^{*}\zeta^{*t}K \\ H_{\zeta v}{}^{*}(1-\zeta^{*t}KN) & H_{\zeta v}{}^{*}\zeta^{*t}K \end{bmatrix} \qquad (2.59)$$

It is pointed out by Kosut and Anderson (1986) that the linear terms $\tilde{x}_L$ are almost identical to the linearized systems analyzed by Rohrs (1982). In the case studied in Kosut and Johnson (1984) and Kosut and Anderson (1986), $\tilde{x}_L$ serves as the input to a nonlinear system (see Figure 2.5). This nonlinear model is used to develop local stability conditions. Their local stability theorem is stated as follows. If there exist finite positive constants g, ϵ and $\delta(\epsilon)$ such that

$$\gamma_{\infty}(F) \leq g < 1/\epsilon \qquad (2.60)$$

$$|\tilde{x}|<\delta(\epsilon) \text{ implies } |f(\tilde{x})|<\epsilon|\tilde{x}| \qquad (2.61)$$

then

$$||\tilde{x}_L||_{\infty} \leq (1-g\epsilon)\delta(\epsilon) \qquad (2.62)$$

implies

$$||\tilde{x}||_{\infty} \leq \delta(\epsilon) \qquad (2.63)$$

where $\gamma_{\infty}(\cdot)$ and $||\cdot||_{\infty}$ denote L_{∞} gain and L_{∞}-norm, respectively. This theorem is based on the linearization theorem of Desoer and Vidyasagar (1975).

As can be seen from (2.60-2.63) the adaptive system, $\tilde{x}$, is stable and confined to a bounded region if the linear system, $\tilde{x}_L$, is bounded and sufficiently small and $F\epsilon L_{\infty}$-stable. However no statement is made in this theorem regarding how these conditions are satisfied. From the definition of the tuned system, it is known that $e^* \epsilon L_\infty$, $\zeta^* \epsilon L_\infty$ and $H_{ev}{}^*$, $H_{\zeta v}{}^* \ \epsilon\ L_\infty$-stable. Therefore from the expressions for $\tilde{x}_L$ and F (2.58,2.59), $\tilde{x}_L \epsilon L_\infty$ and $F \epsilon L_\infty$-stable if and only if

$$\tilde{\theta}_L \ \epsilon\ L_\infty.$$

From (2.58)

$$\tilde{\theta}_L = [I+LM]^{-1}\tilde{\theta}_0 + [I+LM]^{-1}L\zeta^* e^*$$

Therefore

$$s\tilde{\theta}_L + \gamma M\tilde{\theta}_L = s\tilde{\theta}_0 + \gamma\zeta^* e^*$$

or

$$\dot{\tilde{\theta}}_L = -\gamma M\tilde{\theta}_L + \gamma\zeta^* e^* \qquad (2.64)$$

Therefore $\tilde{\theta}_L$ is the solution of the differential equation in (2.64). Kosut and Anderson (1986) select persistent excitation as a mechanism for ensuring stability of (2.64). Consider the system

$$\dot{x} = -\gamma f H f^t x + \gamma w \qquad (2.65)$$

Anderson (1977) shows that if $f \epsilon PE$ and $H(s) \epsilon SPR$, then $(x(0),w) \rightarrow x$ is exponentially stable. This result is used to provide stability of (2.64) by rewriting this equation in a form similar to (2.65), i.e.

$$\dot{x} = -\gamma\zeta^* \bar{H}_{ev}\zeta^{*t} x + \gamma w - Qx \qquad (2.66)$$

where

$$H_{ev}{}^* = \bar{H}_{ev} + \tilde{H}_{ev}$$

$$Q = \gamma(M - \zeta^* \bar{H}_{ev}\zeta^{*t})$$

$$w = \zeta^* e^*$$

$\bar{H}_{ev}$ is referred to as the nominal representation of $H_{ev}{}^*$ and $\tilde{H}_{ev}$ is the deviation caused by modeling error. Intuitively speaking, from comparing (2.66) with (2.65), if $\zeta^* \epsilon PE$, $\hat{H}_{ev} \epsilon SPR$ and Q is sufficiently small, then the system in (2.66) is exponentially stable. These conditions are stated precisely in Kosut and Anderson (1986). Therefore persistent excitation provides boundedness of $\tilde{x}_L$ and stability of F. If $\tilde{x}_L$ is sufficiently small (2.62) then the adaptive system is locally stable.

2.6 Conclusions

One of the key contributions made by Kosut and Johnson (1984) is the development of an error feedback system for the analysis of adaptive control systems. This has enabled the use of standard input-output stability theory (e.g. passivity). An example based on a model reference adaptive controller with a constant gain parameter estimation scheme has been used throughout the development to demonstrate the applicability of this approach. The global stability results of Kosut and Johnson (1984) for continuous-time adaptive systems in the presence of unmodeled dynamics lead to a strictly positive real (SPR) condition on an operator which is a function of the unmodeled dynamics. These results are limited by this SPR condition which is almost surely violated due to typically unmodeled, high frequency dynamics. This conclusion justifies considering local rather than global stability results when unmodeled dynamics are known to be present.

2.7 References

Anderson, B.D.O. (1977). Exponential stability of linear equations arising in adaptive identification. *IEEE Trans. Aut. Contr.*, AC-22, 83-88.

Desoer, C.A. and M. Vidyasagar (1975). *Feedback Systems: Input-Output Properties*, Academic Press, New York.

Doyle, J.C. and G. Stein (1981). Multivariable feedback design: concepts for a classical/modern synthesis. *IEEE Trans. Aut. Contr.*, AC-26, 4-16.

Egardt, B. (1979). *Stability of Adaptive Controllers*, Springer-Verlag, Berlin.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Kosut, R.L. and B.D.O. Anderson (1986). A local stability analysis for a class of adaptive systems. *IEEE Trans. Aut. Contr.*, AC-31, 86-89.

Kosut, R.L. and B. Friedlander (1985). Robust adaptive control: conditions for global stability. *IEEE Trans. Aut. Contr.*, AC-30, 610-624.

Kosut, R.L. and C.R. Johnson, Jr. (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Kreisselmeier, G. and K.S. Narendra (1982). Stable MRAC in the presence of bounded disturbances. *IEEE Trans. Aut. Contr.*, AC-27, 1169-1175.

Martin-Sanchez, J.M. (1976). A new solution to adaptive control. *Proc. IEEE*, 64, 1209-1218.

Narendra, K.S., Y.H. Lin and L.S. Valavani (1980). Stable adaptive controller design, part II: proof of stability. *IEEE Trans. Aut. Contr.*, AC-25, 440-448.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Rohrs, C. (1982). Adaptive control in the presence of unmodeled dynamics. Dept. of Elec. Eng. Comput. Sci., MIT, LIDS Report No. TH-1254.

3. Robustness of Discrete-Time Adaptive Controllers

3.1 Introduction

This chapter is a tutorial/discussion based primarily on the work of R. Ortega, L. Praly and I.D. Landau in the area of robust adaptive control as presented in Ortega *et al* (1985). The major tool for studying the robustness of discrete adaptive control systems is the sector stability theorem (Safanov, 1980). Ortega *et al* (1985) separate the overall adaptive system (i.e. controller plus plant) into two subsystems; one representing the model-plant mismatch and the other representing the parameter adaptation algorithm. This framework was first used by Kosut and Johnson (1984) to study the robustness of continuous-time adaptive control systems. The main contribution of this work by Ortega *et al* (1985) is that it combines Kosut and Johnson's (1984) error system approach with Gawthrop and Lim's (1982) conic sector analysis for discrete-time systems.

3.2 Problem Formulation

3.2.1 Plant and Controller Definition

The actual plant to be controlled may be represented by the equation

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + \xi(k) \tag{3.1}$$

where A and B are polynomials in the backward shift operator

q^{-1}, A is monic, d is the delay and ξ is a bounded output disturbance. The order and coefficients of the polynomials are unknown.

The controller selected for analysis by Ortega *et al* (1985) is an all-zero cancelling scheme with closed loop poles equal to the roots of the polynomial $C_R(q^{-1})$ where the tracking error is

$$e(k) = C_R y(k) - r(k) \tag{3.2}$$

and r(k) is the reference signal. The regulator structure is derived from the predictive control law by equating an estimate of the process output at time k+d to the reference signal at the same instant, i.e.

$$\begin{aligned} r(k+d) &= \hat{S}(k)u(k)+\hat{R}(k)y(k) \\ &= \hat{\theta}(k)^t\phi(k) \end{aligned} \tag{3.3}$$

where $\hat{S}$ and $\hat{R}$ are polynomials functions in q^{-1} of degrees n_s and n_r with time-varying coefficients. The dimension of $\hat{\theta}$ and ϕ is determined by the selected model order. At this point, Ortega *et al* (1985) make an assumption with regards to the existence of a stabilizing parameter set. A vector θ^* is defined as

$$\theta^{*t} = [s_0^*, s_1^* \cdots s_{n_s}^*; r_0^*, r_1^* \cdots r_{n_r}^*] \tag{3.4}$$

and the polynomial C is defined as

$$C = AS^* + q^{-d}R^*B \tag{3.5}$$

Combining (3.5) with (3.1) gives

$$Cy(k) = B\theta^{*t}\phi(k-d) + S^*\xi(k) \tag{3.6}$$

$$Cu(k) = A\theta^{*t}\phi(k) - R^*\xi(k) \tag{3.7}$$

Writing the control law (3.3) with $\hat{\theta}(k)$ replaced by θ^* and inserting into (3.6) gives

$$Cy(k) = Br(k) + S^{*}\xi(k) \tag{3.8}$$

Therefore, BC^{-1} represents the closed loop transfer function of the plant with the fixed parameter (θ^{*}) controller. Ortega *et al* (1985) now assume that there exists a nonempty set θ_{LS} defined as

$$\theta_{LS}=\{\theta^{*}:C(q)\neq 0\ \forall q \text{ such that } |q|>\mu^{1/2}\} \tag{3.9}$$

where $\mu\in(0,1)$ is a scalar. θ_{LS} defines a set of fixed parameter controllers which ensure that the closed loop poles (i.e. roots of C) are stable and lie within a disc of radius $\mu^{1/2}$ where μ is a designer selected parameter which will be defined later. Ortega *et al* (1985) state that if θ_{LS} is empty then the plant cannot be stabilized even when it is perfectly known.

3.2.2 Error Model

Ortega *et al* (1985) define the parameter Ψ as

$$\begin{aligned}\Psi(k) &= (\hat{\theta}(k-d)-\theta^{*})^{t}\phi(k-d) \\ &= \tilde{\theta}(k-d)^{t}\phi(k-d)\end{aligned} \tag{3.10}$$

The tracking error in (3.2) may now be written as

$$\begin{aligned}e(k) &= C_{R}y(k) - r(k) \\ &= C_{R}C^{-1}B\theta^{*t}\phi(k-d)+C_{R}C^{-1}S^{*}\xi(k)-r(k) \quad \text{(from (3.6))}\end{aligned} \tag{3.11}$$

$$\begin{aligned}&= -C_{R}C^{-1}B(r(k)-\theta^{*t}\phi(k-d)) \\ &\qquad +C_{R}C^{-1}Br(k)-r(k)+C_{R}C^{-1}S^{*}\xi(k)\end{aligned} \tag{3.12}$$

$$e(k) = -H_{2}\Psi(k) + e^{*}(k) \tag{3.13}$$

$$\Psi(k) = H_{1}e(k) \tag{3.14}$$

where

$$e^*(k) = (H_2-1)r(k) + C_R C^{-1} S^* \xi(k) \tag{3.15}$$

$$H_2 = C_R C^{-1} B \tag{3.16}$$

Thus from (3.8), $C_R^{-1}H_2=y(k)/r(k)$ represents the transfer function of the process in closed loop with the stabilizing controller. Combining (3.6) and (3.7), the regressor ϕ may be expressed as

$$\phi(k-d) = -W_1\Psi(k) + \phi^*(k-d) \tag{3.17}$$

where

$$\phi^*(k-d) = W_1 r(k) + W_2\xi(k) \tag{3.18}$$

$$W_1=C^{-1}[A\ q^{-1}A \cdots q^{-n_s}A;q^{-d}B \cdots q^{-d-n_r}B]^t \tag{3.19}$$

$$W_2=C^{-1}[-q^{-d}R^* \cdots -q^{-d-n_s}R^*;q^{-d}S^* \cdots q^{-d-n_r}S^*]^t \tag{3.20}$$

If an upper bound on the order of A and B is known, then there exists S^* and R^* such that

$$C = C_R B \tag{3.21}$$

Therefore from (3.16) and (3.21), $H_2=1$ in the matched case.

A schematic of the error model in (3.13) and (3.14) is presented in Figure 3.1. Many similarities exist between this model and the error system developed by Kosut and Johnson (1984) for a continuous adaptive controller. For instance e^* and ϕ^* represent inputs into the feedback system in Figure 3.1 and similar tuned signals act as external inputs into Kosut and Johnson's (1984) system. In both models, the relation H_1 which maps $e \rightarrow \Psi$ contains information on the parameter adaptation algorithm (PAA) and H_2 represents a relation which contains information on the model-plant mismatch (MPM).

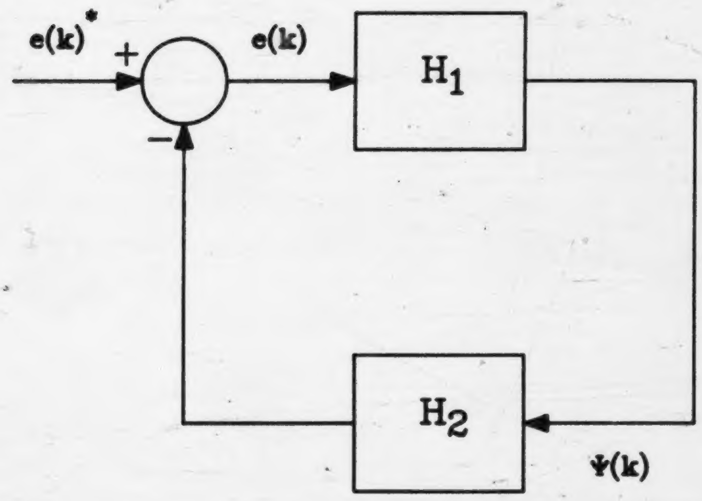
e(k)*
+
e(k)
H1
−
H2
Ψ(k)

Figure 3.1 Error model schematic

3.3 Two Parameter Adaptation Algorithms

The stability technique used by Ortega *et al* (1985) to analyze the error system is conic sector stability theory. Safanov (1980) states that for two operators in a feedback interconnection, such as H_1 and H_2, if one operator is *outside* a sector and the inverse of the other operator is *strictly inside* the sector, then the closed loop system is globally stable.

Gawthrop (1980) derived conic sector properties for a H_1-type operator based on a least squares PAA. However to make use of these results Gawthrop (1980) used an *a priori* boundedness asumption on the regressor. To avoid this assumption in their stability analysis, Ortega *et al* (1985) concluded that normalization of e and ϕ is necessary in order that the PAA sector properties are independent of a boundedness assumption on ϕ. The normalized variables and operators used by Ortega *et al* (1985) are as follows:

$$\phi^n(k-d) = \rho(k)^{-1/2}\phi(k-d) \tag{3.22}$$

$$e^n(k) = \rho(k)^{-1/2}e(k) \tag{3.23}$$

$$\Psi^n(k) = \rho(k)^{-1/2}\Psi(k) \tag{3.24}$$

$$H^n = \rho(k)^{-1/2}H[\rho(k)^{1/2}\cdot] \tag{3.25}$$

The normalization factor ρ will be presented later in more detail.

Ortega *et al* (1985) have ruled out the use of gain decreasing PAA's because an adaptive controller should be able to track variations in the plant behaviour. They have selected two algorithms for further analysis.

1) Constant Gain (CG) PAA

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + f\phi^n(k-d)e^n(k) \qquad (3.26)$$

$f \epsilon R,\ f>0$

2)Regularized Least Squares (RLS) PAA

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + F(k)\phi^n(k-d)e^n(k) \qquad (3.27)$$

$$F(k)=(1-\lambda_0/\lambda_1)\left[F(k-d)-\frac{F(k-d)\phi^n(k-d)\phi^n(k-d)^t F(k-d)}{\lambda+\phi^n(k-d)^t F(k-d)\phi^n(k-d)}\right]+\lambda_0 I \qquad (3.28)$$

where $\lambda_0<\lambda_1$, λ are strictly positive scalars. The eigenvalues of $F(k)$ are contained in the interval $[\lambda_0,\lambda_1]$. Expressions (3.26) or (3.27-3.28) describe the operator $H_1^n : e^n(k) \to \Psi^n(k)$. Ortega *et al* (1985) also consider an exponentially weighted version of $H_1^n(k)$ for the RLS/PAA, i.e. $\bar{H}_1^n : \bar{e}^n(k) \to \bar{\Psi}^n(k)$ where

$$\bar{x}(k) = a^k x(k) \quad a>0. \qquad (3.29)$$

Note that $\bar{H}_1^n = H_1^n$ when $a=1$. The sector properties of these two PAA's are stated below.

1) CG/PAA

$H_1^n + \sigma_{CG}/2$ is passive

for all $\sigma_{CG} \geq f$

2) RLS/PAA

H_1^n is outside the cone$(-1,(1-\sigma_{RLS})^{1/2})$

for all $\sigma_{RLS} \geq \lambda_1/(\lambda+\lambda_1)$

3.4 Stability Analysis

3.4.1 Normalized System

From the input-output properties for $H_1{}^n$ and the sector stability theorem (Safanov, 1980), conditions on $H_2{}^n$ are derived which ensure L_2 stability of the normalized system represented by

$$\Psi^n(k) = H_1{}^n e^n(k) \tag{3.30}$$

$$e^n(k) = -H_2{}^n \Psi^n(k) + e^n(k)^* \tag{3.31}$$

which is just the normalized version of (3.13) and (3.14). For the CG/PAA, if $H_2{}^n$ is strictly inside the cone(C_A,R_A) where

$$(C_A,R_A)=(\sigma_{CG}{}^{-1},\sigma_{CG}{}^{-1})$$

for any $\sigma_{CG} \geq f$, then $e^n(k)$, $\Psi^n(k) \epsilon L_2$ for all $e^n(k)^* \epsilon L_2$. The L_2 and L_∞ stability results for the RLS/PAA are also presented by Ortega *et al* (1985) but are omitted here.

3.4.2 Extension to the Original Error System

At this point Ortega *et al* (1985) present their choice of the normalization factor.

$$\rho(k)=\mu\rho(k-1)+\max(|\phi(k-d)|^2,\bar{\rho}) \tag{3.32}$$

$$\bar{\rho}>0,\ \mu \epsilon (0,1)$$

The problem to be solved is to establish stability of the error system in (3.13) and (3.14) from its normalized counterpart. From the L_2 stability results for the normalized system

$$\lim_{k\to\infty} e^n(k)^2=\lim_{k\to\infty} e(k)^2/\rho(k)=0 \tag{3.33}$$

If the normalization factor had been selected as

$$\rho(k)=1+|\phi(k-d)|^2 \qquad (3.34)$$

then Goodwin *et al*'s (1980) Key Technical Lemma could have been used to show that (3.33) combined with a minimum phase asssumption on the process implies

$$\lim_{k\to\infty} e(k)=0 \qquad (3.35)$$

One of the conditions in the KTL is the uniform boundedness condition, i.e. for a general $\rho(k)$ of the form

$$\rho(k) = b_1(k) + b_2(k)\cdot|\sigma(k)|^2 \qquad (3.36)$$

$b_1(k)$ and $b_2(k)$ must be finite. Since $b_1(k)=b_2(k)=1$ in (3.34) the uniform boundedness condition is satisfied for this normalization factor. However for $\rho(k)$ as selected by Ortega *et al* (1985)

$$b_1(k) = \mu\rho(k-1) \qquad (3.37)$$

In this case the uniform boundedness condition is not satisfied (i.e. $\rho(k)$ is bounded only if ϕ is assumed bounded) and hence the KTL cannot be used to prove that $e(k)\epsilon L_2$.

A procedure for solving this problem of proving stability of the error signals from the stability of the normalized signals has been presented by Ortega *et al* (1985) and is reviewed below.

Ortega *et al* (1985) define exponentially weighted signals as

$$x^{\mu}(k) = \mu^{-k/2}x(k) \qquad (3.38)$$

Summing (3.32) from 0 to N gives

$$\rho(N)=\mu^N\rho(0) + \sum_{k=1}^{N}\mu^{N-k}\max(|\phi(k-d)|^2,\bar{\rho}) \tag{3.39}$$

$$\le\mu^N\rho(0) + \sum_{k=1}^{N}\mu^{N-k}|\phi(k-d)|^2 + \sum_{k=1}^{N}\mu^{N-k}\bar{\rho} \tag{3.40}$$

Multiplying through by μ^{-N} yields

$$\mu^{-N}\rho(N)\le\rho(0) + \sum_{k=1}^{N}\mu^{-k}|\phi(k-d)|^2 + \sum_{k=1}^{N}\mu^{-k}\bar{\rho} \tag{3.41}$$

Using the formula for a geometric progression

$$\sum_{k=1}^{N}\mu^{-k}=(\mu^{-N}-1)\cdot(1-\mu)^{-1}$$

$$<\mu^{-N}\cdot(1-\mu)^{-1} \quad (\text{since } \mu\epsilon(0,1)) \tag{3.42}$$

Combining (3.41) and (3.42)

$$\mu^{-N}\rho(N)\le\rho(0) + \sum_{k=1}^{N}\mu^{-k}|\phi(k-d)|^2+\mu^{-N}\bar{\rho}\cdot(1-\mu)^{-1} \tag{3.43}$$

From comparing (3.43) with Ortega *et al*'s (1985) expression (5.3) it appears that they have defined

$$||\phi^{\mu}(k-d)||_N{}^2 = \sum_{k=1}^{N}\mu^{-k}|\phi(k-d)|^2 \tag{3.44}$$

It may be shown that if $H(q^{-1})$ is a rational function in the backward shift operator and if

$$y(k) = H(q^{-1})\cdot x(k) \tag{3.45}$$

then

$$y^{\mu}(k) = H[(\mu^{1/2}q)^{-1}]\cdot x^{\mu}(k) \tag{3.46}$$

Therefore (3.17)

$$\phi(k-d)=-W_1\Psi(k)+W_1r(k)+W_2\xi(k)$$

implies that

$$\mu^{-k/2}\phi(k-d)=-W_1[(\mu^{1/2}q)^{-1}]\mu^{-k/2}\Psi(k) \tag{3.47}$$

$$+W_1[(\mu^{1/2}q)^{-1}]\mu^{-k/2}r(k)$$

$$+W_2[(\mu^{1/2}q)^{-1}]\mu^{-k/2}\xi(k)$$

Since W_1 and W_2 are stable by assumption then applying the truncated L_2 norm to (3.47) gives

$$||\phi^{\mu}(k-d)||_N\le\gamma_2'[||\Psi^{\mu}(k)||_N+||r^{\mu}(k)||_N]$$

$$+\gamma_2''[||\xi^{\mu}(k)||_N] \tag{3.48}$$

where γ_2' and γ_2'' are the L_2 gains of $W_1[(\mu^{1/2}q)^{-1}]$ and $W_2[(\mu^{1/2}q)^{-1}]$, respectively. From the definition of $\Psi^n(k)$ in (3.24) and (3.43)

$$\Psi^n(N)^2 = \Psi(N)^2 \cdot \rho(N)^{-1} \tag{3.49}$$

$$\geq \mu^{-N}\Psi(N)^2 \cdot \{\rho(0) + ||\phi^\mu(k-d)||_N{}^2 + \mu^{-N}\bar{\rho}'\}^{-1} \tag{3.50}$$

where $\bar{\rho}' = \bar{\rho}(1-\mu)^{-1}$. Using the following identity

$$(a+b)^2 \leq 2(a^2+b^2) \tag{3.51}$$

and combining this with (3.50) gives

$$\begin{aligned}\Psi^n(N)^2 \geq \mu^{-N}\Psi(N)^2 \cdot \\ \{\rho(0) + \mu^{-N}\bar{\rho}' + 4[\gamma_2']^2(||\Psi^\mu(k)||_N{}^2 + ||r^\mu(k)||_N{}^2) \\ + 2[\gamma_2'']^2||\xi^\mu(k)||_N{}^2\}^{-1}\end{aligned} \tag{3.52}$$

Since $\Psi^n(k) \epsilon L_2$ (i.e. $\Psi^n(k) \to 0$) then $\forall\ \delta > 0$, there exists N_0 such that $\forall\ N \geq N_0$

$$\Psi^n(N)^2 \leq \delta \tag{3.53}$$

Therefore from (3.52) and (3.53)

$$\begin{aligned}\mu^{-N}\Psi(N)^2 \leq \delta\{\rho(0) + \mu^{-N}\bar{\rho}' + 4[\gamma_2']^2(||\Psi^\mu(k)||_N{}^2 + ||r^\mu(k)||_N{}^2) \\ + 2[\gamma_2'']^2||\xi^\mu(k)||_N{}^2\}\end{aligned} \tag{3.54}$$

From the definition given in (3.44)

$$\begin{aligned}||\Psi^\mu(k)||_N{}^2 &= \sum_{k=1}^{N}[\Psi^\mu(k)]^2 \\ &= \sum_{k=1}^{N}[\mu^{-k/2}\Psi(k)]^2 \\ &= \sum_{k=1}^{N}\mu^{-k}\Psi(k)^2 \\ &= \sum_{k=1}^{N-1}\mu^{-k}\Psi(k)^2 + \mu^{-N}\Psi(N)^2 \\ &= ||\Psi^\mu(k)||_{N-1}{}^2 + \mu^{-N}\Psi(N)^2\end{aligned} \tag{3.55}$$

Combining (3.54) and (3.55) yields

$$\mu^{-N}\psi(N)^2(1-4\delta[\gamma_2']^2)\leq\delta\rho(0)+\delta\mu^{-N}\tilde{\rho}'+4\delta[\gamma_2']^2||\psi^\mu(k)||_{N-1}^2$$
$$+4\delta[\gamma_2']^2||r^\mu(k)||_N^2$$
$$+2\delta[\gamma_2'']^2||\xi^\mu(k)||_N^2 \tag{3.56}$$

Ortega *et al* (1985) have reexpressed (3.56) as

$$\mu^{-N}\psi(N)^2\leq\delta^2K_1\mu^{-N}+\frac{4\delta[\gamma_2']^2}{1-4\delta[\gamma_2']^2}\sum_{k=N_0}^{N-1}\mu^{-k}\psi(k)^2 \tag{3.57}$$

where they have used the fact that $\rho(0)$, $\tilde{\rho}'$, $r(k)$, $\xi(k)$ and $\{\psi\}_0^{N_0}$ ϵ L_∞ to bound them by $\delta K_1\mu^{-N}$. However this requires the following intermediate step.

$$||r^\mu(k)||_N^2=\sum_{k=1}^{N}\mu^{-k}r(k)^2$$
$$\leq\sup r(k)^2\sum_{k=1}^{N}\mu^{-k}$$
$$\leq\sup r(k)^2\cdot\mu^{-N}(1-\mu)^{-1} \tag{3.58}$$

Ortega *et al* (1985) then apply the Bellman-Gronwall Lemma to (3.57). From Desoer and Vidyasagar (1975)

If $h(k)\geq0$

$$\text{and } u(k)\leq f(k)+\sum_{j=0}^{k}h(j)u(j)\quad\forall k \tag{3.59}$$

and if for some constant h_m, $h(j)\leq h_m$ $\forall j$, then

$$u(k)\leq f(k)+h_m\sum_{j=0}^{k}(1+h_m)^{k-j-1}f(j) \tag{3.60}$$

From comparing (3.57) and (3.59)

$$u(N)=\mu^{-N}\psi(N)^2$$
$$f(N)=\delta^2K_1\mu^{-N}$$
$$h_m=4\delta[\gamma_2']^2\cdot\{1-4\delta[\gamma_2']^2\}^{-1}$$

and substituting into (3.60) gives

$$\mu^{-N}\psi(N)^2\leq\delta^2K_1\mu^{-N} \tag{3.61}$$
$$+\frac{4\delta[\gamma_2']^2}{1-4\delta[\gamma_2']^2}\sum_{k=N_0}^{N-1}(1-4\delta[\gamma_2']^2)^{-(N-k-1)}\delta^2K_1\mu^{-k}$$

which may be rearranged as

$$\Psi(N)^2 \leq \delta^2 K_1 \cdot \{1+4\delta[\gamma_2']^2 \cdot \sum_{k=N_0}^{N-1} \{\mu/(1-4\delta[\gamma_2']^2)\}^{N-k} \qquad (3.62)$$

Ortega *et al* (1985) choose δ such that $1-4\delta[\gamma_2']^2>\mu$, which makes the term inside the summation less than unity and hence the series convergent. Therefore it may be concluded that $\Psi(k)\epsilon L_\infty$.

Ortega *et al* (1985) use a corollary of this result to prove boundedness of the normalization factor $\rho(k)$. Since $r(k)$, $\xi(k)\epsilon L_\infty$ and W_1 and W_2 are stable by assumption then from (3.17), $\Psi(k)\epsilon L_\infty$ implies that $\phi(k)\epsilon L_\infty$ and consequently from (3.32) $\rho(k)\epsilon L_\infty$.

Ortega *et al* (1985) then introduce an important lemma which relates the conic restrictions on $H_2{}^n$ to similar restrictions on H_2. Consider the operator $H:x(k)\rightarrow y(k)$. If $H[(\mu^{1/2}q)^{-1}]$ is inside the cone(C,R), then $H^n(=\rho(k)^{-1/2}H[\rho(k)^{1/2}\cdot])$ with $\rho(k)$ as defined in (3.32) is inside the same cone(C,R).

At this point Ortega *et al* (1985) are able to state their final L_2 stability result. Consider the adaptive controller defined by (3.3) with parameters updated according to (3.26) applied to the process in (3.1). If for given n_s, n_r and μ, and the set θ_{LS} in (3.9) is nonempty, and

(i) $H_2[(\mu^{1/2}q)^{-1}]$ is strictly inside the cone(C_A,R_A)

(ii) $r(k)$, $\xi(k)\epsilon L_\infty$ such that $e(k)^*\epsilon L_2$, then

$\Psi(k)$, $e(k)\epsilon L_2$ and $\phi(k)\epsilon L_\infty$.

Condition (i) ensures the stability of the normalized error system. Since $e(k)^{*} \in L_2$, this implies that $e^{n}(k)^{*} \in L_2$ and hence $\Psi^{n}(k) \in L_2$. As shown above $\Psi^{n}(k) \in L_2$ implies $\Psi(k) \in L_\infty$ which in turn implies $\rho(k) \in L_\infty$. Therefore $\rho(k)$ qualifies as a multiplier in Figure 3.2. From multiplier theory (Desoer and Vidyasagar, 1975) L_2 stability of the error system depicted in Figure 3.2 implies L_2 stability of the original error system (Figure 3.1).

3.5 Conclusions

Ortega *et al* (1985) have developed global stability results for a discrete-time adaptive controller. This work is a combination of the error system approach of Kosut and Johnson (1984) with the conic sector ideas of Gawthrop and Lim (1982). The key modification to the adaptive control algorithm is the concept of normalization used in the parameter estimation scheme.

The global stability results lead to a conic sector condition on the operator H_2 which corresponds to the SPR condition imposed on its continuous-time counterpart. Ortega *et al* (1985) point out that H_2 always has relative degree zero because the delay is assumed known. Ortega *et al* (1985) state further that this implies that for all stably invertible processes the Nyquist locus of H_2 starts and ends on the same side of the imaginary axis. These comments suggest that the conic condition which arises in the discrete-time case may not be as restrictive as the SPR condition for the continuous-time case when unmodeled dynamics are present.

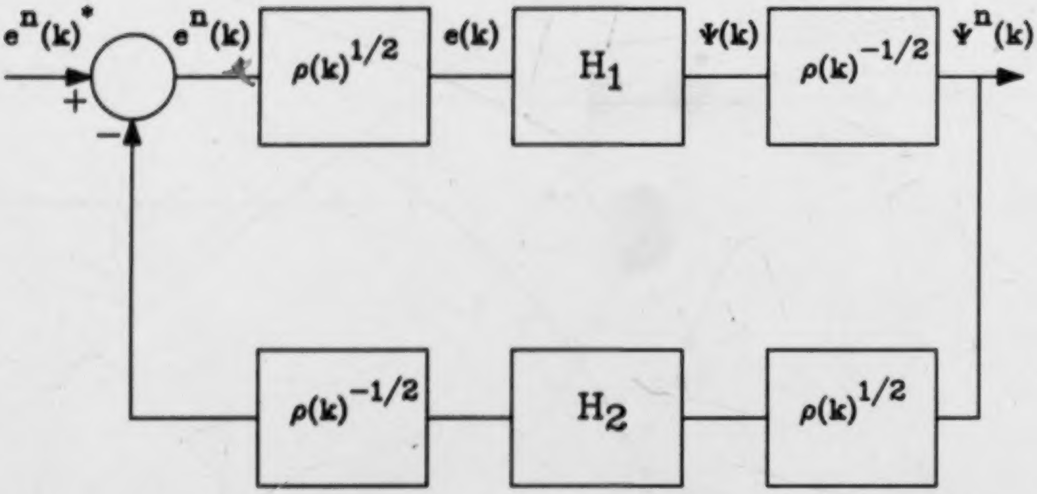
$e^n(k)^*$
$e^n(k)$
$\rho(k)^{1/2}$
$e(k)$
H_1
$\Psi(k)$
$\rho(k)^{-1/2}$
$\Psi^n(k)$
+
−
$\rho(k)^{-1/2}$
H_2
$\rho(k)^{1/2}$

Figure 3.2 Normalized error model schematic

3.6 References

Desoer, C.A. and M. Vidyasagar (1975). *Feedback Systems: Input-Output Properties*, Academic Press, New York.

Gawthrop, P.J. (1980). On the stability and convergence of a self-tuning controller. *Int. J. Contr.*, 31, 973-998.

Gawthrop, P.J. and K.W. Lim (1982). Robustness of self-tuning controllers. *IEE Proc.*, 129, 21-29.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Kosut, R.L. and C.R. Johnson, Jr. (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Safanov, M.G. (1980). *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, MA.

4. Robust Design of Adaptive Control Systems

4.1 Introduction

Rohrs (1982) was one of the first researchers to study the robustness of adaptive controllers. Rohrs performed an extensive investigation of several control schemes available in the literature for both continuous-time and discrete-time systems. All of the algorithms that Rohrs examined require the assumption that an upper bound on the order of the linear plant in known and is used in the formulation of the control law. Rohrs pointed out that this assumption is almost always violated in practice because real plants contain high frequency dynamics which are not, in general, included in the model used for control. Through simulation using a third order plant with second order, high frequency dynamics, and a first order model, Rohrs demonstrated that certain reference signals and/or output disturbances may result in an unstable control system. His primary conclusion from this work was that the adaptive algorithms he considered cannot be used in practice with any confidence because instability will probably occur due to the presence of unmodeled dynamics.

Rohrs (1982) work resulted in two key publications. In Rohrs *et al* (1984) a linearization analysis of an adaptive control system was used to demonstrate analytically some design guidelines for discrete-time adaptive controllers in

---

A version of this chapter has been accepted for publication. Cluett, W.R., S.L. Shah and D.G. Fisher (1986). Automatica.

the presence of unmodeled dynamics. Rohrs *et al* (1985a) demonstrated that two infinite-gain operators exist in continuous-time adaptive control systems and that sinusoidal reference inputs at specific frequencies and output disturbances at any frequency can cause the loop gain to increase and the control system to eventually become unstable.

In this chapter Rohrs' *discrete-time* results (Rohrs, 1982; Rohrs *et al*, 1984) are examined and extended. In section 4.2 the linearization analysis presented by Rohrs *et al* (1984) is reviewed. A comment on Rohrs' results is presented in section 4.3, where it is demonstrated that the simulation examples studied by Rohrs are unstable even when no unmodeled dynamics are present due to a nonminimum phase zero introduced by fast sampling. It is also shown using simulation that the adaptive controller analyzed by Rohrs does in fact have some robustness in the presence of unmodeled dynamics when the minimum phase assumption is satisfied and the sampling interval is selected sufficiently large. In section 4.4 the conic sector stability results of Ortega *et al* (1985) are used to develop a design approach for discrete adaptive control systems which provides a quantitative measure of the effect of the design guidelines discussed by Rohrs *et al* (1984) on stability in the presence of unmodeled plant dynamics.

4.2 Linearization Analysis

4.2.1 Rohrs' Error System

Rohrs *et al* (1984) used projection algorithm II (Goodwin *et al*, 1980) to demonstrate their linearization analysis. The true plant is represented by

$$A(q^{-1})y(k) = q^{-d} \cdot g_p \cdot B(q^{-1})u(k) \tag{4.1}$$

and the reference model is given by

$$A_m(q^{-1})y_m(k) = q^{-d} \cdot g_m \cdot B_m(q^{-1})r(k) \tag{4.2}$$

where A, B, A_m and B_m are polynomials in the backward shift operator q^{-1}. Rohrs defined auxiliary variables as delayed versions of the plant input and output, i.e.

$$w_{yi}(k) = q^{-i}y(k) \quad i=0,1,\cdots,n-1 \tag{4.3}$$

$$w_{ui}(k) = q^{-(i+1)}u(k) \quad i=0,1,\cdots,n-2 \tag{4.4}$$

$$w_r(k) = [g_m B_m/A_m] \cdot r(k) \tag{4.5}$$

The control law is given by

$$u(k) = K(k)^t w(k) \tag{4.6}$$

where K(k) is a vector of time-varying controller parameters and w(k) is a composite vector of the auxiliary variables.

$$K(k) = [k_r(k)\ K_u(k)^t\ K_y(k)^t]^t \tag{4.7}$$

$$w(k) = [w_r(k)\ w_u(k)^t\ w_y(k)^t]^t \tag{4.8}$$

The notation presented in (4.1)-(4.8) is consistent throughout Rohrs (1982) for all algorithms studied. However it does not match the notation originally used by Goodwin *et al* (1980) for the projection algorithm II. For instance, Goodwin *et al* (1980) express the control law in (4.6) as

$$u(k) = \hat{\theta}(k)^t \phi(k) \tag{4.9}$$

where

$$\phi(k)=[-y(k)\cdots-y(k-n+1),$$
$$-u(k-1)\cdots-u(k-m-d+1),y_m(k+d)]^t \qquad (4.10)$$

where n and m are the respective orders of A and B in (4.1). Rohrs *et al* (1984) do not state the dimension of the A and B polynomials in (4.1) but if it is assumed that their dimension n in (4.3) and (4.4) is equivalent to Goodwin *et al*'s (1980) dimension n of their A polynomial, then (4.8) may be expanded and compared to (4.10).

$$w(k)=[y_m(k+d),u(k-1),\cdots u(k-n+1),$$
$$y(k),\cdots y(k-n+1)]^t \qquad (4.11)$$

The negative signs in (4.10) may be absorbed into $\hat{\theta}(k)$ in (4.9) and the arrangement of the elements in (4.10) and (4.11) may be changed to make the vectors look similar. However there is a discrepancy between the order of $w_u(k)$ in (4.11) and the number of 'u' terms in (4.10), i.e. Rohrs has included (n-1) 'u' terms and Goodwin has (m+d-1) 'u' terms in his ϕ vector. In general these two quantities are not equal and therefore the definition of w_u in (4.4) should be modified accordingly.

The parameter adjustment scheme for K(k) in (4.6) was stated by Rohrs *et al* (1984) as

$$K(k) = K(0) - \frac{\gamma}{1-q^{-d}}\left[\frac{w(k-d)e(k)}{1+w(k-d)^t w(k-d)}\right] \qquad (4.12)$$

where

$$e(k) = y(k) - y_m(k) \qquad (4.13)$$

and γ, the adaptive gain in (4.12), is chosen such that $\gamma g_p<2$. Goodwin *et al* (1980) state the same adaptive algorithm as

$$\hat{\theta}(k) = \hat{\theta}(k-d)-\hat{\beta}_0^{-1}[\frac{\phi(k-d)e(k)}{1+\phi(k-d)^t\phi(k-d)}] \qquad (4.14)$$

where $\hat{\beta}_0$ is chosen such that $0<\beta_0/\hat{\beta}_0<2$. By comparison with (4.12) γ should have the additional condition of γg_p being strictly positive. Also g_p in (4.1) should not be interpreted as a process gain but is equivalent to β_0, the first coefficient in the B' polynomial of the following alternate process representation to (4.1)

$$A(q^{-1})y(k) = q^{-d}B'(q^{-1})u(k) \qquad (4.15)$$

where $B'=g_pB$. Rohrs *et al* (1984) then introduced a fixed parameter control law with K(k) in (4.6) replaced by K^* where

$$u(k) = q^{-1}K_u^*u(k)+K_y^*y(k)+k_r^*y_m(k+d) \qquad (4.16)$$

$$K_u^* = k_{u0}^*q^{-1}+\cdots+k_{u(n-2)}^*q^{-(n-1)} \qquad (4.17)$$

$$K_y^* = k_{y0}^*+\cdots+k_{y(n-1)}^*q^{-(n-1)} \qquad (4.18)$$

If (4.17) and (4.18) are substituted into (4.16) then

$$\begin{aligned}u(k) = {} & k_{u0}^*u(k-2)+\cdots+k_{u(n-2)}^*u(k-n) \\ & +k_{y0}^*y(k)+\cdots+k_{y(n-1)}^*y(k-n+1) \\ & +k_r^*y_m(k+d) \qquad (4.19)\end{aligned}$$

From comparing (4.19) with the time-varying form in (4.6) and (4.11), it is seen that the two control laws do not contain the same 'u' terms. If the q^{-1} multiplying the first term on the right-hand side of (4.16) is omitted then this discrepancy is removed. If the fixed parameter control law in (4.16) is used with the plant in (4.1), then a tuned closed loop transfer function, $q^{-d}g^*B^*/A^*$, may be obtained

$$\begin{aligned} y(k) &= Pu(k) \qquad (P=q^{-d}g_pB/A) \\ &= P[K_u^*(1-K_u^*)^{-1}\cdot(K_y^*y(k)+k_r^*y_m(k+d)) \\ &\qquad + K_y^*y(k)+k_r^*y_m(k+d)] \end{aligned} \tag{4.20}$$

and collecting terms

$$\begin{aligned} & y(k)\cdot(1-PK_u^*(1-K_u^*)^{-1}K_y^*-PK_y^*) \\ &\quad = PK_u^*(1-K_u^*)^{-1}k_r^*y_m(k+d)+Pk_r^*y_m(k+d) \end{aligned} \tag{4.21}$$

which may be rewritten as

$$\begin{aligned} \frac{y(k)}{y_m(k+d)} &= \frac{Pk_r^*}{1-K_u^*-PK_y^*} \\ &= \frac{q^{-d}g_pk_r^*B}{A(1-K_u^*)-q^{-d}g_pBK_y^*} \\ &= \frac{q^{-d}g^*B^*}{A^*} \end{aligned} \tag{4.22}$$

Rohrs *et al* (1984) pointed out that if the plant order and model order are equal then it is possible to choose K^* such that

$$y(k)/y_m(k+d) = q^{-d} \tag{4.23}$$

and the controlled plant output will match the model output ($e(k)=y(k)-y_m(k)=0$). When the model order is less than the actual plant order, the error $e(k)$ is equivalent to a nonlinear expression. The error system used by Rohrs *et al* (1984) is derived below. Define

$$\tilde{K}(k) = K(k)-K^* \tag{4.24}$$

Substituting (4.24) into (4.1) gives

$$y(k) = P[\tilde{K}(k)^t w(k)+K^{*t}w(k)] \tag{4.25}$$

$$= P\tilde{K}(k)^t w(k)+P[K_u{}^*u(k)+K_y{}^*y(k)+k_r{}^*y_m(k+d)] \tag{4.26}$$

$$= (1-K_u{}^*-PK_y{}^*)^{-1}\cdot[P\tilde{K}(k)^t w(k)+Pk_r{}^*y_m(k+d)] \tag{4.27}$$

Using (4.2), (4.22) and (4.27)

$$y(k) = \frac{1}{k_r{}^*}\cdot\frac{q^{-d}g^*B^*}{A^*}\cdot\tilde{K}(k)^t w(k) + \frac{q^{-d}g^*B^*}{A^*}\cdot\frac{g_mB_m}{A_m}\cdot r(k) \tag{4.28}$$

Therefore e(k) may be expressed as

$$e(k) = y(k)-y_m(k)$$

$$= \frac{q^{-d}g^*B^*}{k_r{}^*A^*}\cdot[\tilde{K}(k)^t w(k)] + \frac{q^{-d}g_mB_m}{A_m}\cdot[\frac{g^*B^*}{A^*}-1]\cdot r(k) \tag{4.29}$$

Equation (4.29) and the parameter update scheme in (4.12) define a nonlinear error system as shown in Figure 4.1 which describes the overall closed loop plant. It is at this point that Rohrs *et al* (1984) decided to linearize their error system. Rohrs defined $\tilde{w}(k)$ as

$$\tilde{w}(k) = w(k)-w^* \tag{4.30}$$

and then linearized (4.29) about w^* and K^* by assuming that the plant input, output and reference signal, and the controller parameters are close to their desired final values so that $\tilde{w}$ and $\tilde{K}$ are small. Replacing w(k-d) in (4.12) by w^* gives

$$\tilde{K}(k) = \frac{-\gamma}{1-q^{-d}}\cdot[\frac{w^*e(k)}{1+w^{*t}w^*}] \tag{4.31}$$

where $K(0)=K^*$.

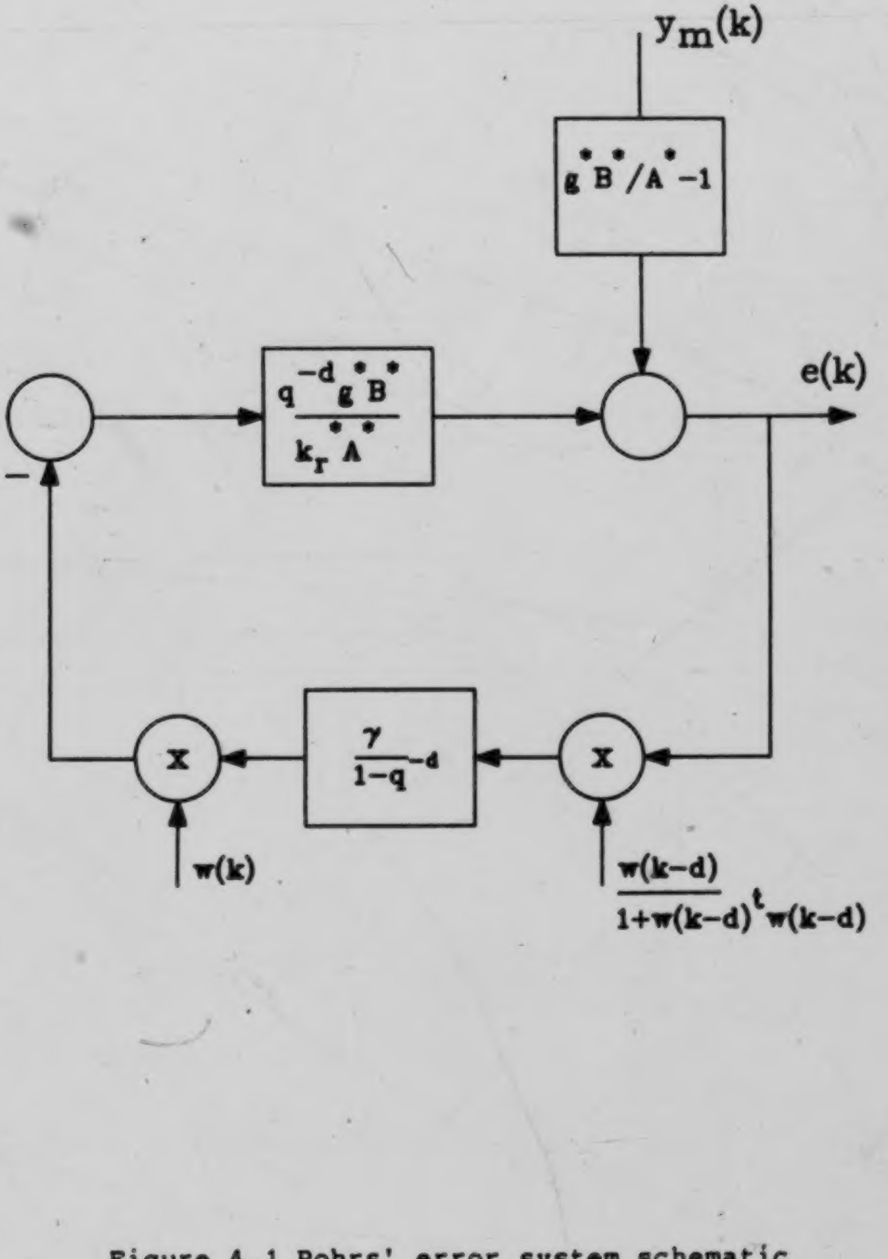
$y_m(k)$
g^*B^*/A^*-1
$q^{-d}\frac{g^*B^*}{k_r A^*}$
$e(k)$
$-$
$\frac{\gamma}{1-q^{-d}}$
x
x
$w(k)$
$\frac{w(k-d)}{1+w(k-d)^t w(k-d)}$

Figure 4.1 Rohrs' error system schematic

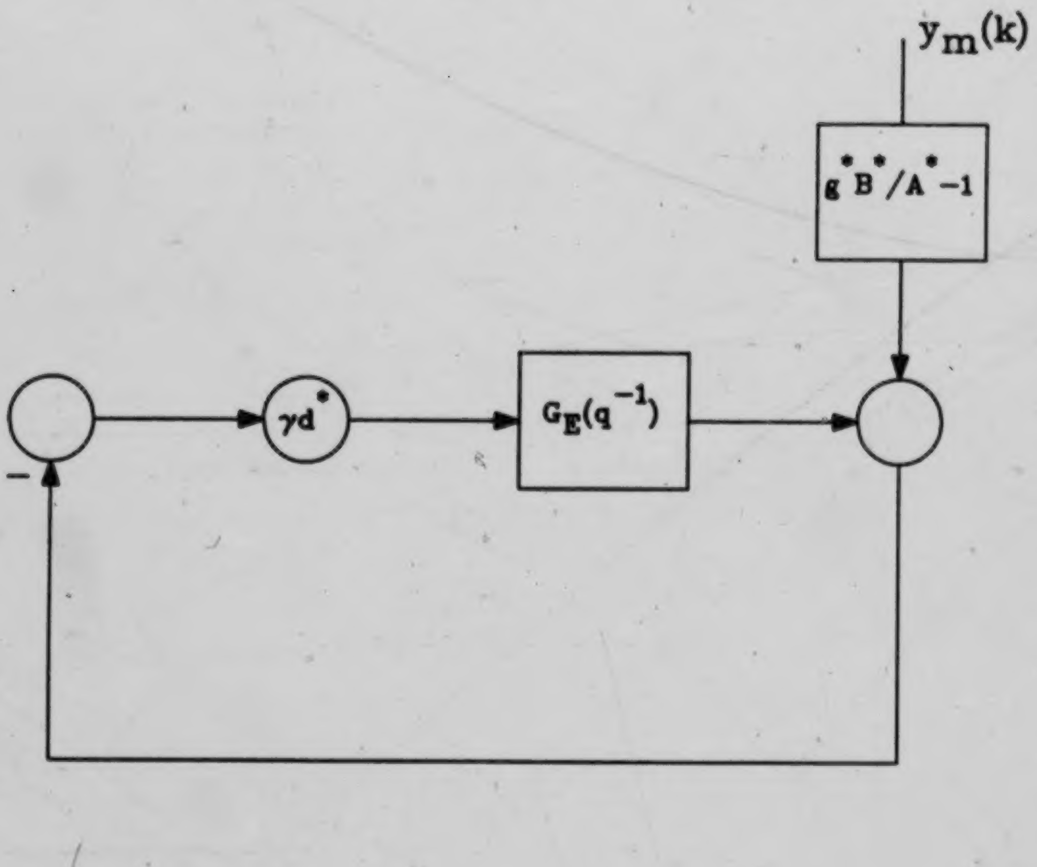
$y_m(k)$
g^*B^*/A^*-1
γd^*
$G_E(q^{-1})$
$-$

Figure 4.2 Rohrs' linearized error system schematic

The error system may then be approximated by the linearized system in Figure 4.2 with

$$d^* = w^{*t}w^*/(1+w^{*t}w^*) \tag{4.32}$$

$$G_E(q^{-1}) = \frac{q^{-d}g^*B^*}{(1-q^{-d})k_r{}^*A^*} \tag{4.33}$$

The analysis on this linearized error system is performed by first choosing a desired system $q^{-d}g^*B^*/A^*$ and $k_r{}^*$ and then analyzing the behaviour of this system at the desired operating point by performing a root locus for the error sytem using γd^* as the variable gain parameter.

4.2.2 Rohrs' Design Guidelines

Rohrs *et al* (1984) first did a simple analysis on a system with no unmodeled dynamics. As pointed out earlier it is possible to choose the desired system such that (4.23) is obtained. From Goodwin *et al* (1980) a reasonable choice for $k_r{}^*$ in the matched case is

$$k_r{}^* = g_p{}^{-1} \tag{4.34}$$

which produces a linearized error sytem with

$$G_E(q^{-1}) = g_p q^{-d}/(1-q^{-d}) \tag{4.35}$$

The closed loop transfer function for this system with $d=1$ is

$$\gamma d^* g_p/(q-1+\gamma d^* g_p) \tag{4.36}$$

Therefore the γd^* root locus of (4.36) produces limits on γd^* to ensure stability of

$$0 < \gamma d^* g_p < 2 \tag{4.37}$$

Since $0 \le d^* < 1$ from inspection of (4.32), then (4.37) corresponds exactly to the condition required by Goodwin *et al* (1980) for global stability (i.e. $0<\gamma g_p<2$). Rohrs then did an example of the analysis on a system with unmodeled dynamics. The example consists of a third order plant which includes second order, high frequency dynamics.

$$y(t) = \frac{2}{s+1}\cdot\frac{229}{s^2+30s+229}\cdot u(t) \tag{4.38}$$

The reference model for this scheme was chosen by Rohrs as

$$y_m(t) = [3/(s+3)]\cdot r(t) \tag{4.39}$$

Rohrs used the standard technique of hold equivalence to obtain discrete-time representations for (4.38) and (4.39). Rohrs *et al* (1984) did most of their analysis at a sampling period of T_s=.04s which leads to the following plant description

$$\frac{y(k)}{u(k)} = \frac{.00361q^{-1}(1+.196q^{-1})(1+2.763q^{-1})}{(1-.961q^{-1})[(1-.547q^{-1})^2+(.044q^{-1})^2]} \tag{4.40}$$

$$y_m(k) = [.12q^{-1}/(1-.88q^{-1})]\cdot r(k) \tag{4.41}$$

Rohrs assumed a first order process model for the controller design. This mismatch between the assumed order (1st) and the true plant order (3rd) defines the unmodeled dynamics. The first order model assumption produces a control law of the form in (4.6) with

$$\begin{aligned} K(k) &= [k_r(k) \quad k_{y0}(k)]^t \\ w(k) &= [y_m(k+1) \quad -y(k)]^t \end{aligned} \tag{4.42}$$

Rohrs refers to $q^{-d}g^*B^*/A^*$ in (4.22) as the nominally controlled system (NCS). For the plant sampled at T_s=.04s, the NCS is given by

$$\begin{aligned} NCS={}&q^{-1}(.00361)k_r^*(1+.196q^{-1})(1+2.763q^{-1})\cdot \\ &\{(1-.961q^{-1})[(1-.547q^{-1})^2+(.44q^{-1})^2] \\ &-q^{-1}(.00361)(1+.196q^{-1})(1+2.763q^{-1})k_{y0}^*\}^{-1} \end{aligned} \qquad (4.43)$$

Rohrs performed a root locus on (4.43) using k_{y0}^* as the variable gain. Note that if Rohrs had selected a model order greater than unity, then multivariable root loci plots would have had to have been analyzed in order to determine a vector of K^* values which produce a stable NCS (i.e. a first order model → $K_u^*=0$, $K_y^*=k_{y0}^*$). The set of tuned parameters for the NCS in (4.43) selected by Rohrs *et al* (1984) is

$$k_{y0}^*=0.8,\ k_r^*=1.32 \qquad (4.44)$$

Rohrs *et al* (1984) gave no indication how the value for k_r^* was chosen. Certainly k_r^* does not appear in the characteristic equation of (4.43). However from (4.22) it would make sense to have the d.c. gain of the NCS equal to unity. This appears to be how Rohrs selected the value for k_r^*.

Rohrs then performed the γd^* root locus of the error system in Figure 4.2 with

$$G_E(q^{-1})=\frac{q^{-1}(.0046)(1+.196q^{-1})(1+2.763q^{-1})}{(1-q^{-1})(1.32)(1-.82q^{-1})(1-.79q^{-1})(1-.45q^{-1})} \qquad (4.45)$$

Rohrs showed that this linearized error system is stable for $\gamma d^*<0.35$. Rohrs did a simulation generated with the parameters initialized at the tuned values given in (4.44) and with

$$\gamma=0.2,\ r=10.0$$

The $w(k)$ for this case is given in (4.42). For a constant reference sequence and assuming that the output converges to the setpoint, the w^* for this case may be taken as

$$w^* = [10 \quad 10]^t$$

Therefore d^* in (4.32) is equal to 200/201=.995 and $\gamma d^*=.199$ which satisfies the root locus criterion on γd^* for stability. The simulation for this case was stable.

Rohrs then performed a root locus of the same error system with γd^* fixed at 0.94 and $k_{y0}{}^*$ the variable. Rohrs showed that there is no value for $k_{y0}{}^*$ which will stabilize the system at $\gamma d^*=0.94$. Based on this line of argument Rohrs states that the first design is illustrated.

> *In order to maintain stability in the presence of unmodeled dynamics, it is necessary that the adaptive gain, γ, of the system be kept small and the adaptation proceed slowly* (Rohrs *et al*, 1984).

Rohrs also remarked that a value for γ around 1.0, which he states is the limit for any chance of stability in the above-mentioned analysis, is an order of magnitude smaller than the gain which would be allowed using the ideal guidelines and ignoring the high frequency unmodeled dynamics. For instance, if the plant is assumed to be given by

$$G(s) = 2/(s+1) \qquad (4.46)$$

instead of (4.38) then for $T_s=.04s$, the corresponding discrete representation is

$$G(q^{-1}) = .078q^{-1}/(1-.961q^{-1}) \qquad (4.47)$$

and $\gamma=g_p{}^{-1}=12.8$ is approximately an order of magnitude greater than unity.

The next point that Rohrs *et al* (1984) illustrated was the importance of the NCS. In order that the NCS exactly match the reference model ($e(k)\rightarrow 0$), the NCS must converge to a deadbeat controller (i.e. $NCS\rightarrow q^{-d}$). For this to occur, all of the poles of the NCS must be moved to the origin. However Rohrs illustrated via his root locus arguments that this cannot occur for any value of k_{y0}^* in (4.43) and therefore concluded that in order to solve this problem the following guideline should be used.

> *Design the nominal control loop so that it is robust and that approximate model matching can be easily attained even in the presence of unmodeled dynamics* (Rohrs *et al*, 1984).

The third design parameter that Rohrs analyzed was the effect of sampling time on stability in the presence of unmodeled dynamics. For instance if the sample period is increased to $T_s=.4s$, Rohrs stated the equivalent discrete-time plant representation of (4.38) as

$$y(k) = \frac{q^{-1}(.629)(1+.0399q^{-1})(1+.0048q^{-1})}{(1-.67q^{-1})[(1-.0017q^{-1})^2+(.0018q^{-1})^2]}\cdot u(k) \qquad (4.48)$$

However this plant was incorrectly discretized by Rohrs and the correct representation is given by

$$y(k) = \frac{q^{-1}(.467)(1+.399q^{-1})(1+.0048q^{-1})}{(1-.67q^{-1})[(1-.0017q^{-1})^2+(.0018q^{-1})^2]}\cdot u(k) \qquad (4.49)$$

Rohrs argued, based on the representation in (4.48), that since the poles and zeros of the unmodeled dynamics almost cancel in the slowly sampled system, then the unmodeled dynamics no longer have a destabilizing effect. Therefore larger adaptive gains, closer to the value allowed if no unmodeled dynamics were present, may be used. Rohrs stated that the third guideline is then verified.

> *Sample the system slowly enough to remove the effects of unmodeled dynamics* (Rohrs *et al*, 1984).

For the correct representation in (4.49) it can be seen that one of the zeros has not migrated to the origin ($q=-.399$) and therefore complete cancellation does not occur between the extra poles and zeros which result from the unmodeled dynamics. Hence the slower sampling does not simplify the control problem to the extent suggested by Rohrs *et al* (1984).

4.3 A Comment on Rohrs' Results

Rohrs' (1982) work in adaptive control demonstrated that instabilities can occur when certain algorithms are implemented in the presence of unmodeled dynamics. His results suggest that the applicability of the adaptive control algorithms studied is extremely limited because the assumptions necessary for stability are violated in practice. The objective of this section is to present an alternate interpretation of Rohrs' *discrete-time* results, i.e. the instabilities are not entirely due to the presence of unmodeled dynamics. It will be demonstrated that the nonminimum phase zero which arises due to fast sampling of

Rohrs' example causes stability problems even when no unmodeled dynamics are present. It is also shown that the adaptive controllers analyzed by Rohrs have some robustness properties in the presence of unmodeled dynamics when the minimum phase assumption is satisfied and the sampling time is selected properly.

As pointed out in the previous section, Rohrs selected projection algorithm II of Goodwin *et al* (1980) as one of the algorithms for his robustness analysis. This model reference adaptive scheme was proven to be globally stable by Goodwin *et al* (1980) for plants described by

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) \tag{4.50}$$

$$A(q^{-1})=1+a_1q^{-1}+\cdots+a_nq^{-n}$$

$$B(q^{-1})=b_0+b_1q^{-1}+\cdots+b_mq^{-m}$$

which satisfy the following assumptions:

(a) d is known

(b) an upper bound for n and m is known

(c) B(q) has all roots inside the unit circle

The control law is given by (4.9) and (4.10) with the parameter estimates updated according to (4.14).

The objective in Rohrs (1982) and Rohrs *et al* (1984) was to analyze the stability of such 'stable' algorithms in the presence of unmodeled dynamics, i.e. violation of assumption (b). The example used by Rohrs for his simulation work was the third order plant in (4.38). Most of the analysis was done with a sampling period of T_s=.04s which leads to the plant representation in (4.40). In (4.40) a nonminimum phase zero is present due to the short sampling period as explained by Astrom *et al* (1984). In fact this unstable zero does not disappear until T_s>0.2s. (In Figure 4.3 a locus of the zero locations of the plant in (4.38) as a function of sampling time from T_s=.04s to .4s is presented.) Rohrs assumed a first order process model for the controller design in order to violate assumption (b). The first order model assumption produces a control law of the form given by (4.9) and (4.10) with n=1, m=0 and d=1. (Note that this plant also violates assumption (c) due to the nonminimum phase zero.)

Rohrs (1982) and Rohrs *et al* (1984) simulated the process (4.40) plus the reduced order controller with various reference inputs r(t). The results presented here are for the same adaptive controller with no unmodeled dynamics. For the third order plant in (4.40) the control law *without* unmodeled dynamics is given by (4.9) and (4.10) with n=3, m=2 and d=1. Figure 4.4 can be compared directly with Rohrs *et al* (1984) Figure 10 (reproduced in Figure 4.5). The somewhat unexpected result is that the system with no unmodeled dynamics (Figure 4.4) is unstable whereas Rohrs' example (Figure 4.5), which has unmodeled dynamics, is stable. The unstable response in Figure 4.6 for the system with no unmodeled dynamics is directly comparable to the unstable response in Rohrs *et al*'s (1984) Figure 11 (reproduced in Figure 4.7) which is for the same example with unmodeled dynamics.

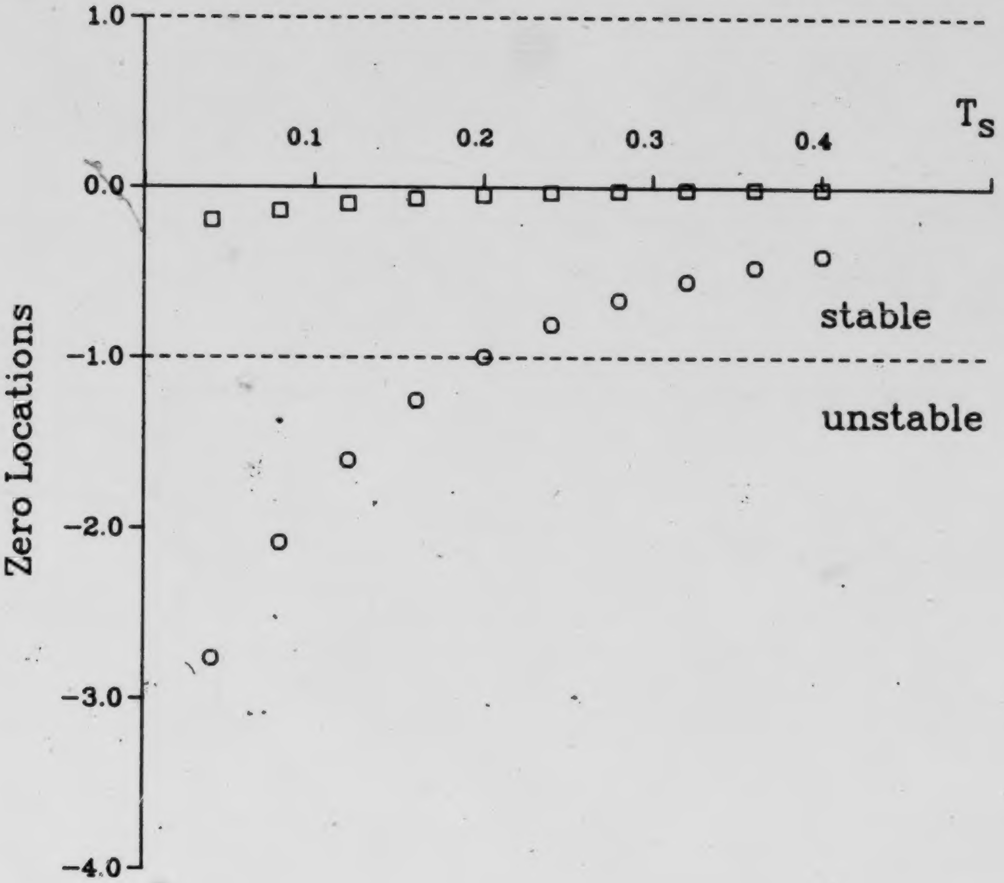

Figure 4.3 Discrete plant zero locations versus sample time

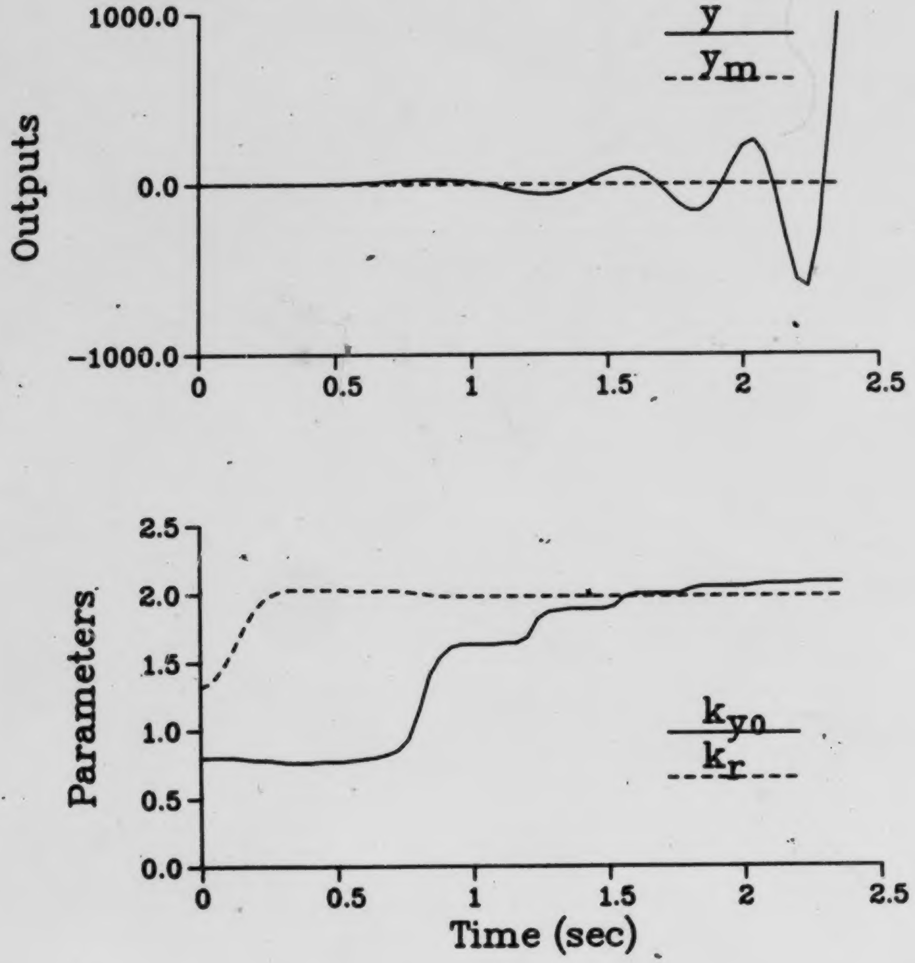

Figure 4.4 Simulation with no unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$

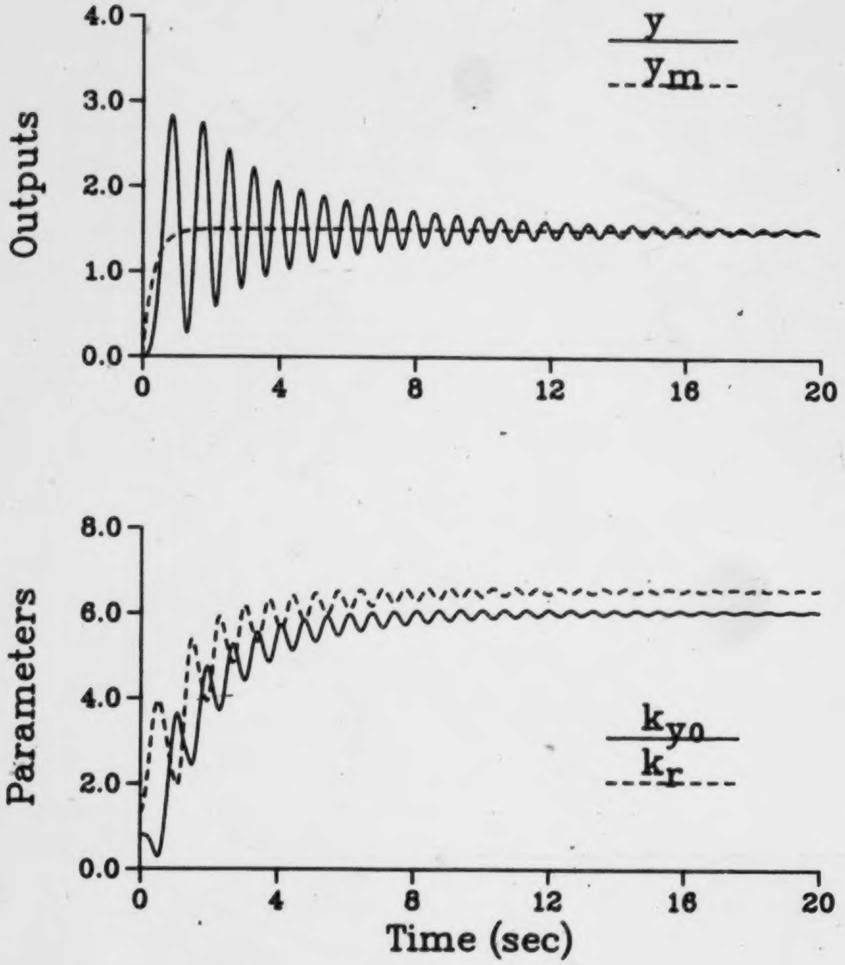

Figure 4.5 Simulation with unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$

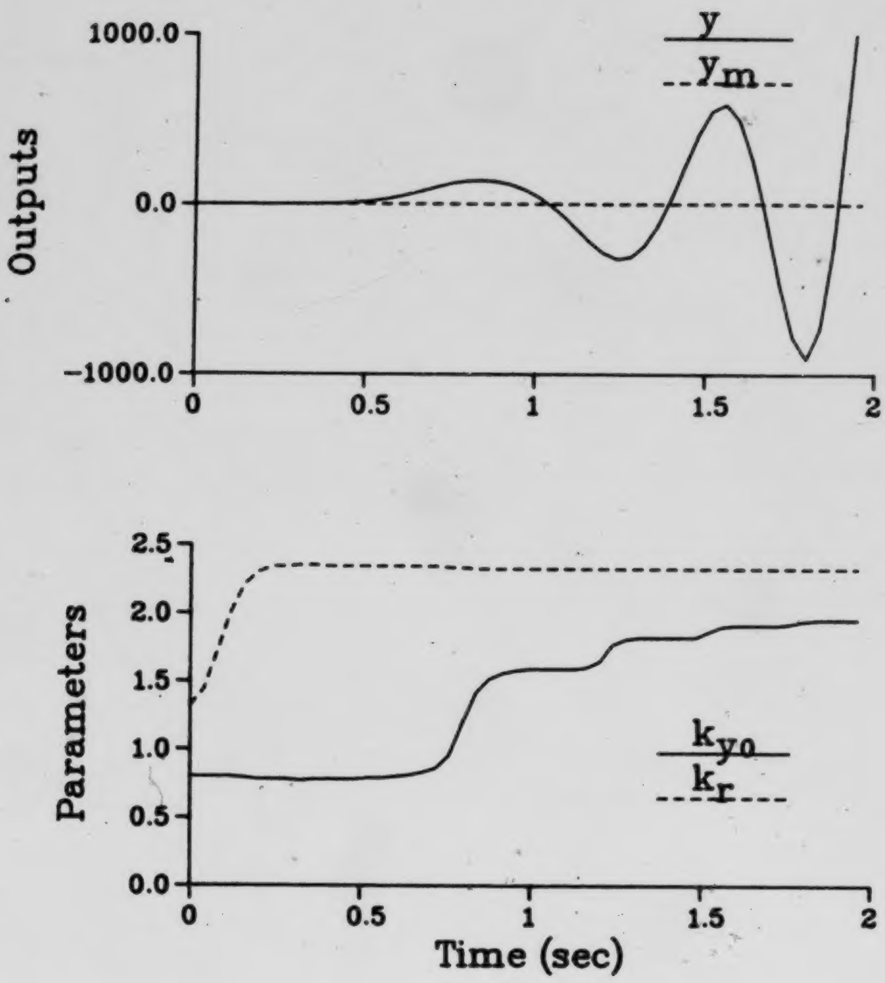

Figure 4.6 Simulation with no unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$

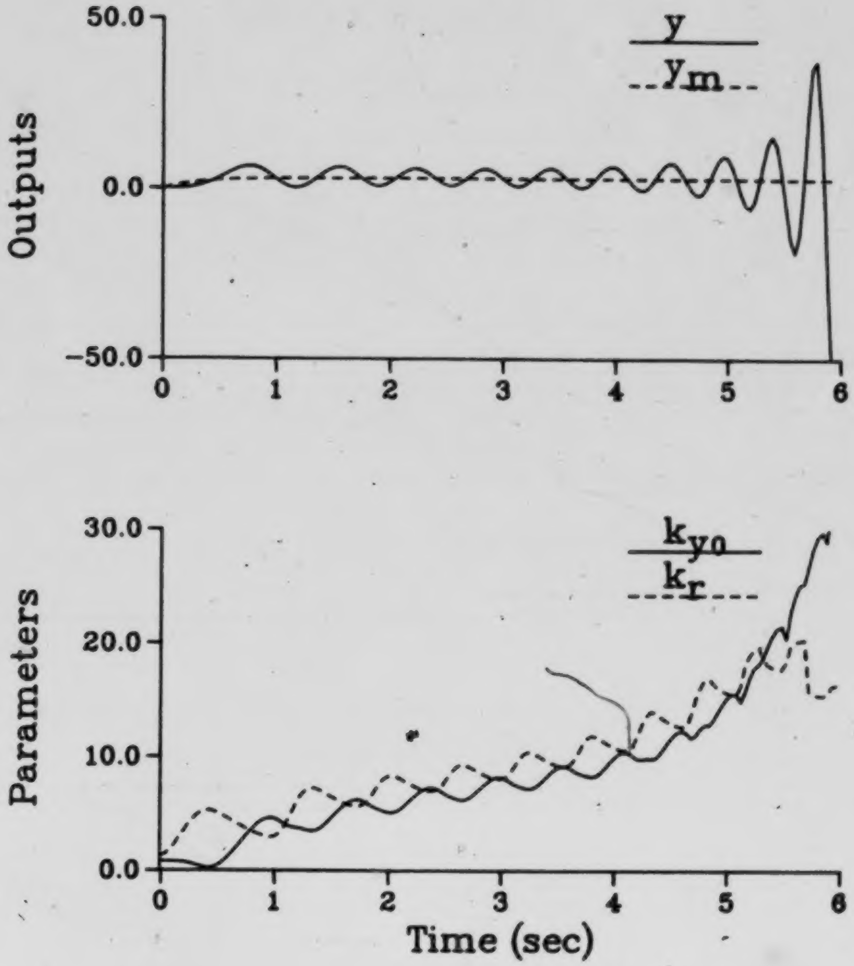

Figure 4.7 Simulation with unmodeled dynamics: r(t)=3.1 and

$\gamma = 1.0$

Without any unmodeled dynamics the only assumption which is violated is that of a stable-inverse plant. This suggests that the instability which was observed in Figure 4.7 by Rohrs is not entirely due to the presence of unmodeled dynamics but also due to the unstable zero which results from fast sampling.

This suggestion is further reinforced by examining the sinusoidal reference input results from Rohrs (1982). Figure 4.8 is a similar simulation result to the one shown in Figure 5-28 of Rohrs (1982). For this case

$r(t) = 1.0 + 4.5\sin(13.5t)$, $\gamma = .1$

(This sinusoidal frequency of the reference input was reported incorrectly in Rohrs (1982) as: $\sin(13.5t/.04)$.) Figure 4.8 shows that the plant becomes unstable when unmodeled dynamics are present. Figure 4.9 is the result under the same conditions except that the unmodeled dynamics have been removed by using the full order control law. This system also becomes unstable solely due to the presence of the unstable zero.

Rohrs also examined the plant at a slower sampling rate (T_s=.4s). The discrete-time representation with this sampling period is given in (4.49). It is clear from (4.49) that this plant has a stable-inverse and therefore does not violate assumption (c). If the full order control law is used to control the discrete process in (4.49), all three of the plant assumptions are satisfied and hence global stability is ensured. Figures 4.10, 4.11 and 4.12 are

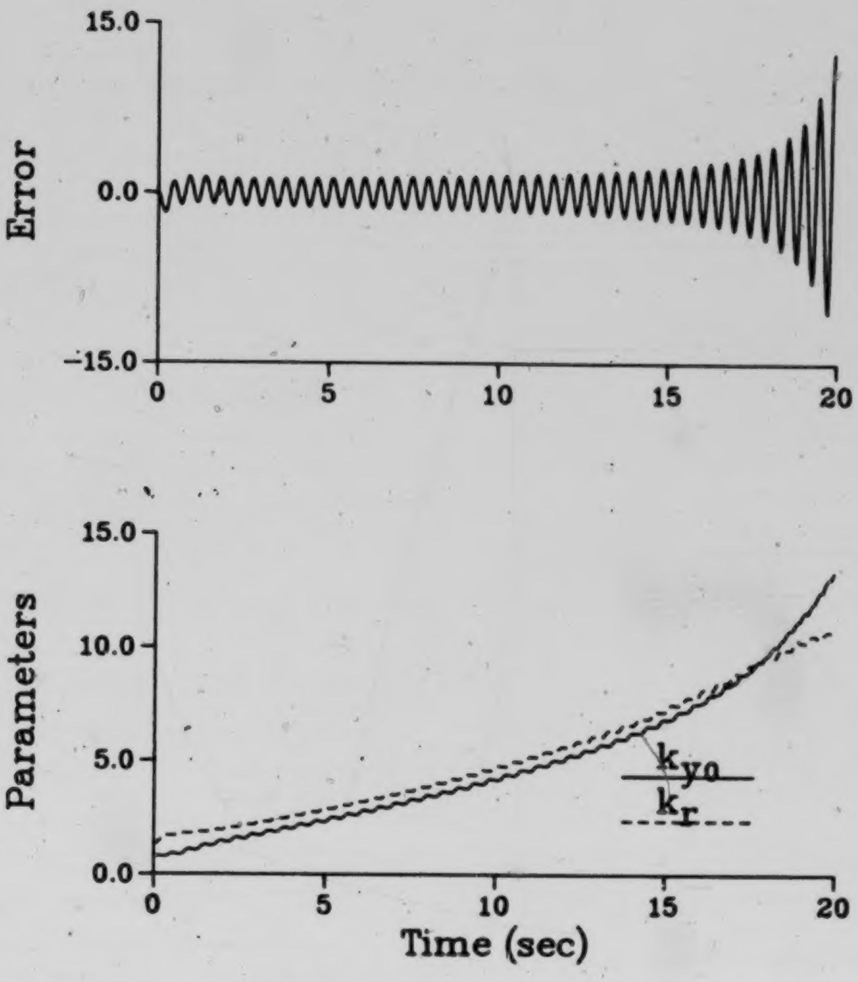

Figure 4.8 Simulation with unmodeled dynamics:

$r(t)=1.0+4.5\sin(13.5t)$ and $\gamma=0.1$

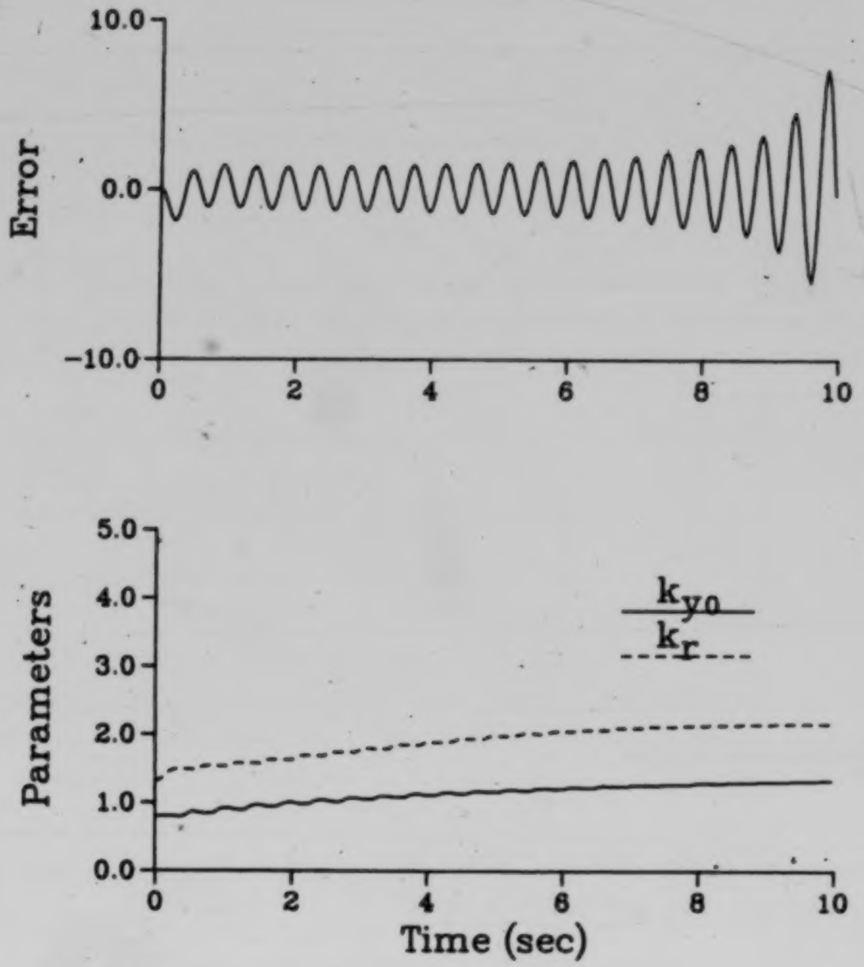

Figure 4.9 Simulation with no unmodeled dynamics:

$r(t)=1.0+4.5\sin(13.5t)$ and $\gamma=0.1$

simulation results with the different types of inputs examined earlier for the process sampled at the slower rate plus the third order control law. All three results confirm that the system is stable and the tracking error converges to zero.

Figures 4.13, 4.14 and 4.15 show the results with the slower sampling rate and unmodeled dynamics present (i.e. the reduced order control law). It is interesting to discover that these three results are *stable*. Therefore despite the presence of unmodeled dynamics the reduced order controller is able to stabilize the discrete-time, stable-inverse process in (4.49). (The reference signal frequency (ω=6.75 rad/s) was chosen as the frequency at which the open-loop discrete plant in (4.49) provides 180^0 phase shift. Other reference signals with frequencies ranging from ω=1.0 to 34.0 rad/s were examined and these results were also stable.)

Rohrs' conclusions in Rohrs (1982) state that the discrete adaptive controllers studied become unstable in the presence of unmodeled dynamics and constant or sinusoidal reference inputs. These conclusions are based almost entirely on simulation results with the discrete-time plant representation in (4.40). In his conclusions, Rohrs seems to have overlooked the fact that the plant violates the stable-inverse assumption. The results in this section clearly point out the instabilities that occur when this assumption is violated with no unmodeled dynamics.

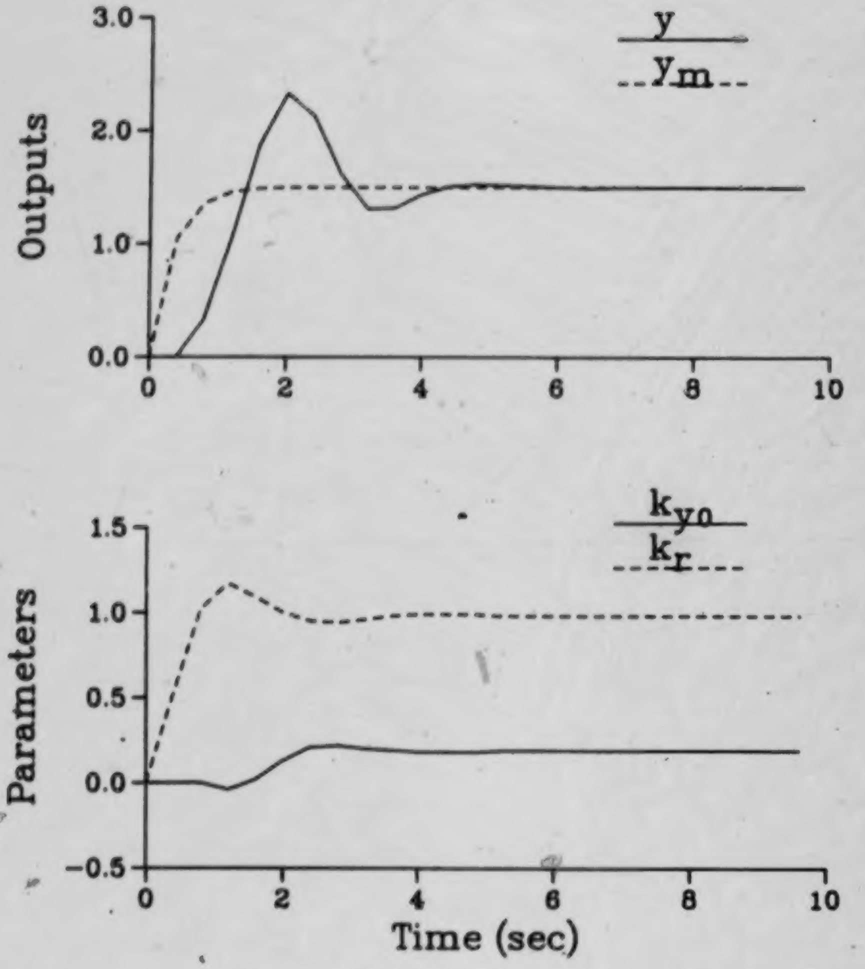

Figure 4.10 Simulation with no unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$

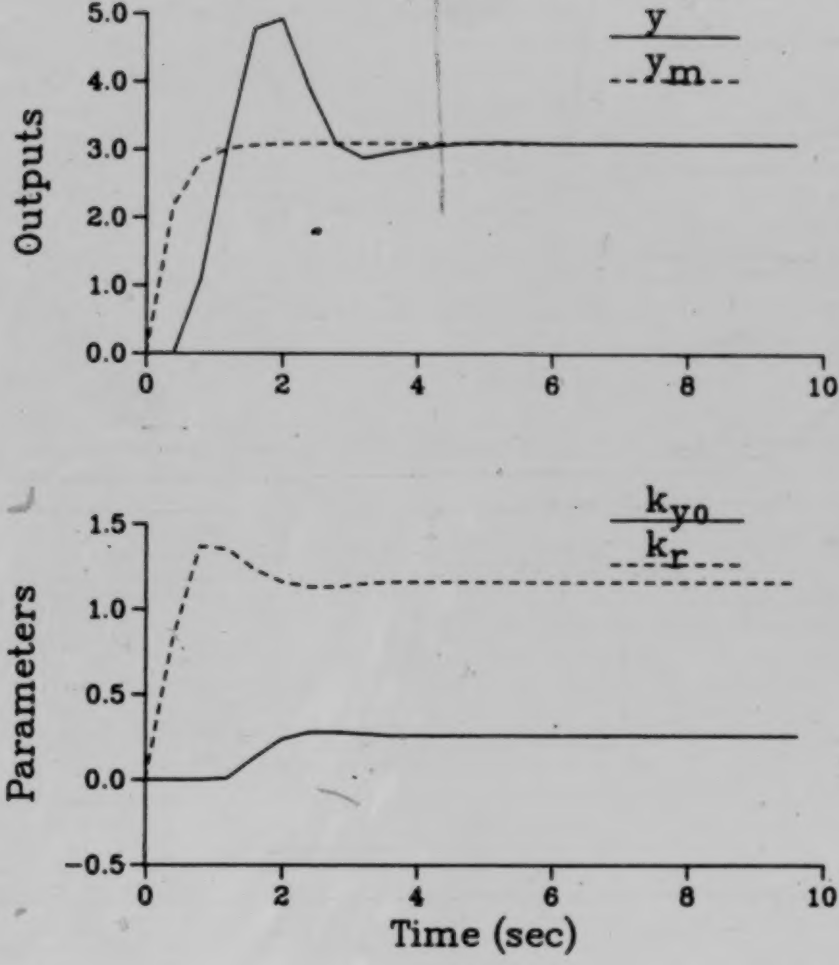

Figure 4.11 Simulation with no unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$

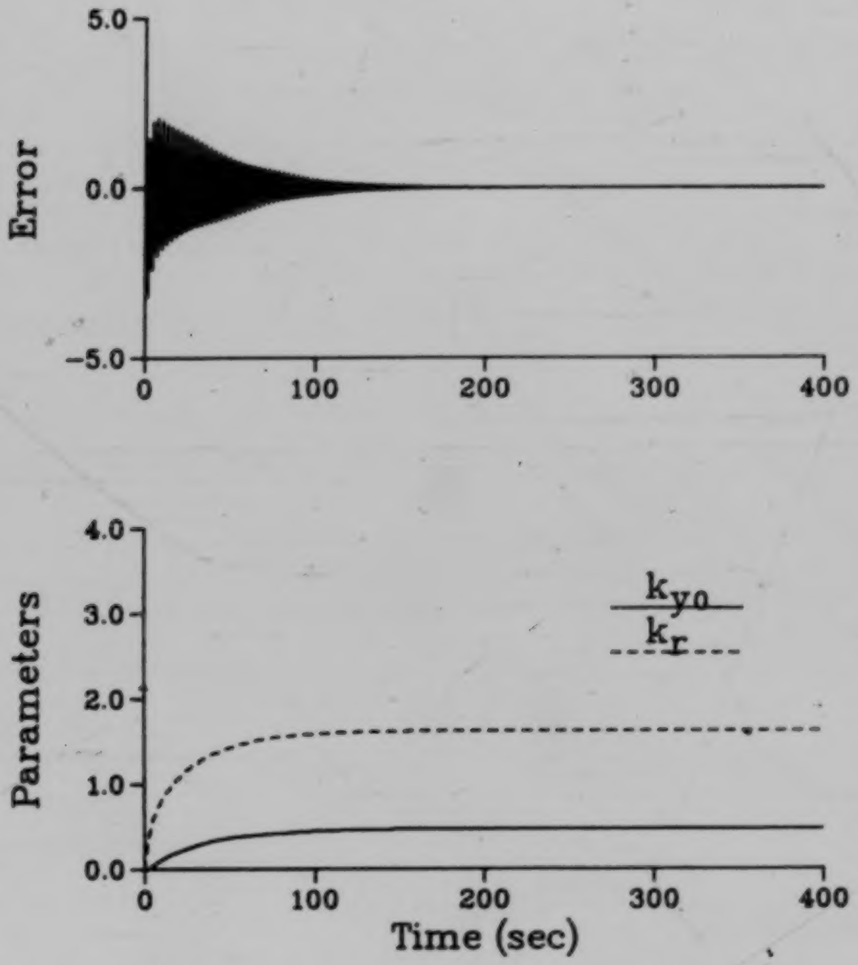

Figure 4.12 Simulation with no unmodeled dynamics:

$r(t)=1.0+4.5\sin(6.75t)$ and $\gamma=0.1$

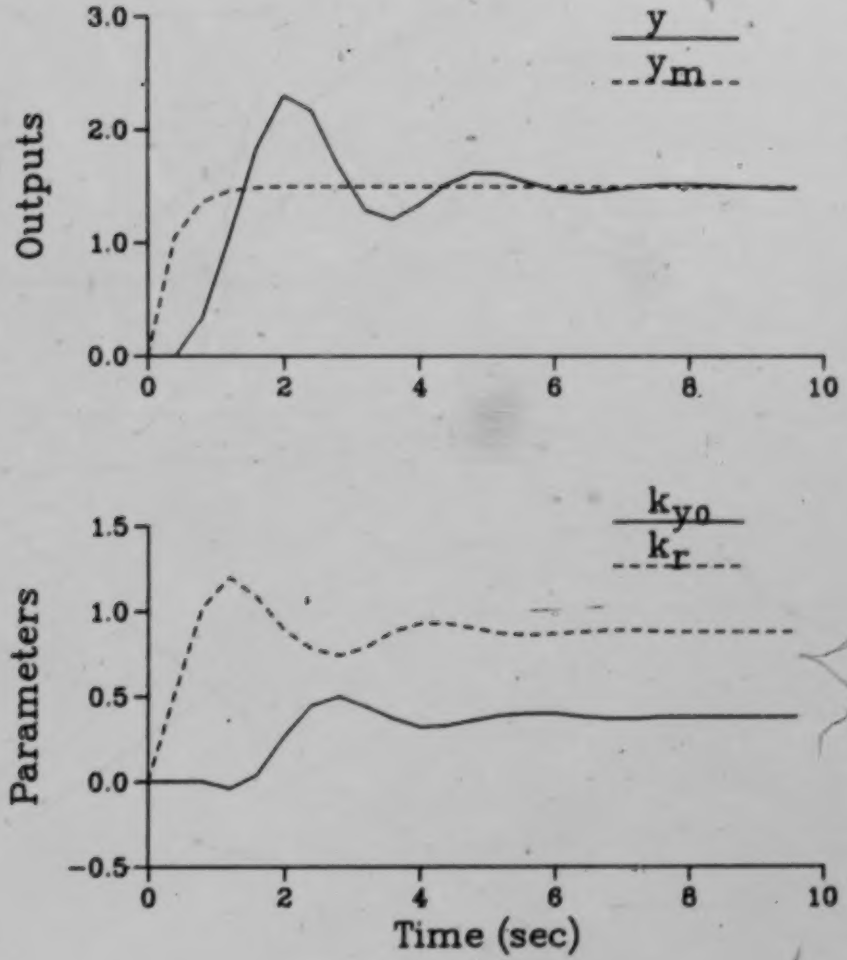

Figure 4.13 Simulation with unmodeled dynamics: $r(t)=1.5$ and $\gamma=1.0$

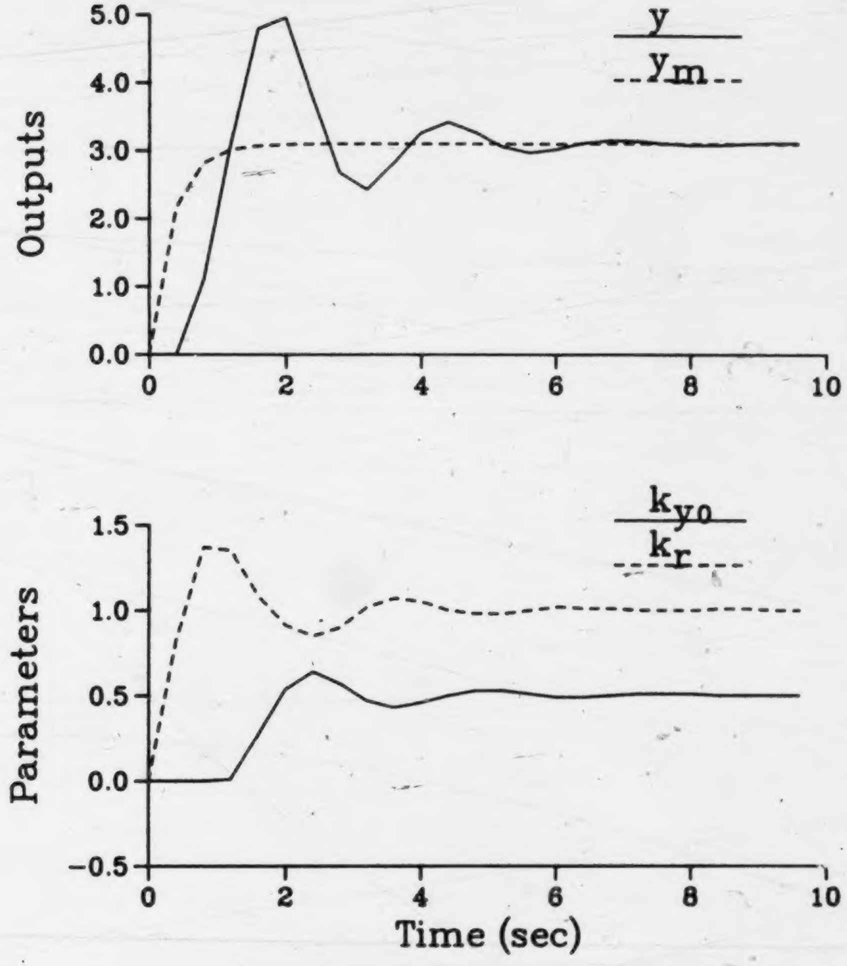

Figure 4.14 Simulation with unmodeled dynamics: $r(t)=3.1$ and $\gamma=1.0$

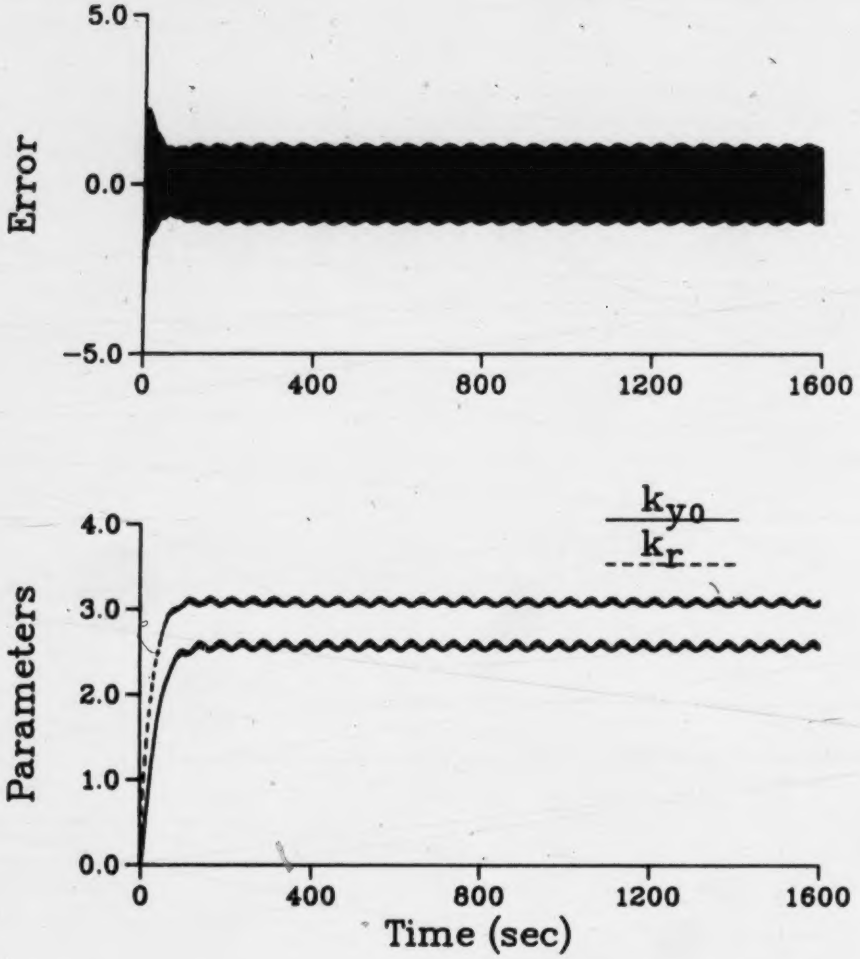

Figure 4.15 Simulation with unmodeled dynamics:

$r(t)=1.0+4.5\sin(6.75t)$ and $\gamma=0.1$

Therefore, any conclusions made about the robustness of the adaptive control algorithms analyzed in Rohrs (1982) should take into account the results presented here.

Rohrs presented some results in Rohrs (1982) with the slower sampled discrete plant in (4.49). However, only the responses with sinusoidal additive output disturbances are examined. Rohrs suggested that since the unmodeled poles and zeros coincide, their effect is negligible and the system behaves as if no unmodeled dynamics are present. According to Rohrs (1982) this is why any problems associated with constant and sinusoidal reference signals are alleviated by slow sampling. As stated in the previous section, this conclusion was based on an incorrect plant representation. From the correct representation in (4.49) it is clear that this simplification does not occur to the extent suggested by Rohrs. Therefore, the plant in (4.49) does contain unmodeled dynamics if the process is assumed to be first order. The results in this section indicate that the adaptive control algorithm of Goodwin *et al* (1980) does have some robustness properties in the presence of unmodeled dynamics with constant and sinusoidal reference inputs. The choice of a slower sampling rate (T_s=.4s) investigated here is not an entirely unreasonable one and is within guidelines proposed by other authors (Astrom and Wittenmark, 1984). More recently, Rohrs *et al* (1985b) have presented results with respect to slow sampling rates which confirm the observations reported here.

Rohrs' (1982) results did show that stability problems associated with the output disturbances are not alleviated by the slower sampling. It appears that Rohrs overlooked algorithms available in the literature at the time (e.g. Egardt, 1979) which handle bounded disturbances using dead-zones or parameter projection. Perhaps it is algorithms such as these which require further attention and analysis when disturbances are present.

4.4 A Design Approach for Discrete Adaptive Control Systems

Rohrs (1982) demonstrated that instabilities can result when adaptive control systems are implemented with unmodeled dynamics present. This work was followed up by Rohrs *et al* (1984) who presented some guidelines for the design of discrete adaptive controllers which improve stability in the presence of unmodeled dynamics. These guidelines were extracted from a linearization analysis of the inherently nonlinear adaptive control problem. This approach enabled the use of linear stability theory (i.e. root locus analysis) to analyze the problem. However, this linearization required some strong assumptions regarding convergence and closeness of the parameters and input/output signals to some desired values.

Until recently little rigorous theory existed in the literature which dealt with the robustness of discrete-time adaptive controllers. Lim (1982), and Ortega *et al* (1985) were among the first researchers to examine the robustness of these controllers in the presence of unmodeled dynamics. Lim's (1982) results are technically incomplete because they use an *a priori* boundedness assumption on the regressor vector. The problem of handling unmodeled dynamics via dead-zones was not considered solvable because of the requirement of *a priori* boundedness of the unmodeled terms (Martin-Sanchez *et al*, 1984). However, recent results using dead-zones plus normalization have allowed various researchers (Cluett *et al*, 1986; Kreisselmeier and Anderson, 1986), to rigorously prove global stability of an adaptive system in the presence of unmodeled dynamics.

In this section a design approach for discrete adaptive control systems is presented which provides a quantitative measure of the effect of design alternatives such as (i) adaptive gain, (ii) model order, and (iii) sampling rate, on stability in the presence of unmodeled dynamics. The proposed method, based on the conic conditions developed by Ortega *et al* (1985), is illustrated using Rohrs' (1982) benchmark example. The results in this section demonstrate that the design guidelines presented by Rohrs *et al* (1984) do indeed have a firm theoretical basis.

The sector condition for stability in the presence of unmodeled dynamics may be interpreted as a strictly positive real (SPR) type of condition for a particular transfer function, H_2, which is a function of the unmodeled dynamics. Many authors (e.g. Kosut and Johnson, 1984) have expressed the opinion that in the *continuous-time* case, the SPR condition is difficult to satisfy and therefore the goal of global stability in adaptive control is unrealistic. However, in the *discrete-time* case this section shows that by making appropriate design tradeoffs it is often possible to satisfy the conic sector condition and hence ensure global stability.

4.4.1 A Robust Stability Result

Ortega *et al* (1985) separated the overall adaptive system (i.e. controller plus plant) into two subsystems; one representing the model-plant mismatch (MPM) and the other representing the parameter adaptation algorithm (PAA). This separation was also used by Kosut and Johnson (1984) for a continuous-time adaptive controller. To effect this separation, the adaptive control system was transformed into an error feedback system which permitted the application of Safanov's (1980) sector stability theorem.

The actual plant to be controlled may be represented by the equation

$$A(q^{-1})y(k)=q^{-d}B(q^{-1})\cdot(1+\tilde{G}(q^{-1}))u(k) \tag{4.51}$$

where A and B are polynomials in the backward shift operator q^{-1} and $\tilde{G}$ is a rational function in q^{-1} which represents the model uncertainty (Doyle and Stein, 1981). The controller analyzed by Ortega *et al* (1985) is an all-zero cancelling scheme with closed loop poles equal to the roots of the polynomial $C_R(q^{-1})$ where the tracking error is

$$e(k)=C_R y(k)-r(k) \tag{4.52}$$

and r(k) is the reference signal. The regulator structure is derived from the predictive control law by equating an estimate of the process output at time k+d to the reference signal at the same instant, i.e.

$$r(k+d)=\hat{S}(k)u(k)+\hat{R}(k)y(k)=\hat{\theta}(k)^t\phi(k) \qquad (4.53)$$

where $\hat{S}$ and $\hat{R}$ are polynomial functions in q^{-1} of degrees n_s and n_r with time-varying coefficients. The dimension of $\hat{\theta}$ and ϕ is determined by the selected model order.

Ortega *et al* (1985) presented two algorithms for updating the control law coefficients. The constant gain parameter adaptation algorithm (CG/PAA) is restated below and will be used for the analysis in this section.

$$\hat{\theta}(k)=\hat{\theta}(k-d)+f\phi(k-d)e(k)/\rho(k) \quad f\epsilon R,\ f>0 \qquad (4.54)$$

$$\rho(k)=\mu\rho(k-1)+\max(|\phi(k-d)|^2,\bar{\rho}) \qquad (4.55)$$

$$\bar{\rho}>0,\ \mu\epsilon(0,1)$$

where $\rho(k)$ is referred to as the normalization factor. The CG/PAA differs from the well-known projection algorithm in which

$$\rho(k)=c+|\phi(k-d)|^2 \quad c>0 \qquad (4.56)$$

and from the stochastic approximation algorithm, where $\mu=1$ in (4.55).

To derive the necessary error system, Ortega *et al* (1985) defined the parameter $\Psi(k)$ as

$$\Psi(k)=(\hat{\theta}(k-d)-\theta^*)^t\phi(k-d)=\tilde{\theta}(k-d)^t\phi(k-d) \qquad (4.57)$$

where θ^* is a vector of stabilizing parameters. Ortega *et al* (1985) assumed that a fixed parameter controller of the same structure as in (4.53) with $\hat{\theta}(k)$ replaced by θ^* ensures a stable closed-loop system. Kosut and Johnson (1984) used a similar assumption for the continuous case. The tracking error in (4.52) may now be expressed as

$$\begin{aligned} &e(k)=-H_2\Psi(k)+e(k)^* \\ &e(k)^*=(H_2-1)r(k) \\ &\Psi(k)=H_1e(k) \end{aligned} \tag{4.58}$$

where H_1 denotes a relation defined by the PAA and H_2 is a relation which contains information on the MPM. A schematic of this error model is given in Figure 4.16.

Two key elements regarding the allowable amount of unmodeled dynamics need to be extracted from the result of Ortega *et al* (1985). First, what is the exact form of H_2 in (4.58), and second, what are the restrictions on H_2 to guarantee global stability?

Ortega *et al* (1985) noted that there exist polynomials S^* and R^* of orders n_s and n_r which verify the identity

$$C_RB=AS^*+q^{-d}R^*B \tag{4.59}$$

If S^* and R^* are chosen to form θ^*, the vector of stabilizing parameters introduced in (4.57), then H_2 may be expressed as

$$\begin{aligned} &H_2=(1+\tilde{G})/(1+\tilde{G}T) \\ &T=q^{-d}R^*/C_R \end{aligned} \tag{4.60}$$

where $\tilde{G}$ is the uncertainty in the plant representation (4.51).

The other requirement is some restrictions on H_2. From the L_2 result given by Ortega *et al* (1985), sufficient conditions for stability are as follows. If for a given n_s,

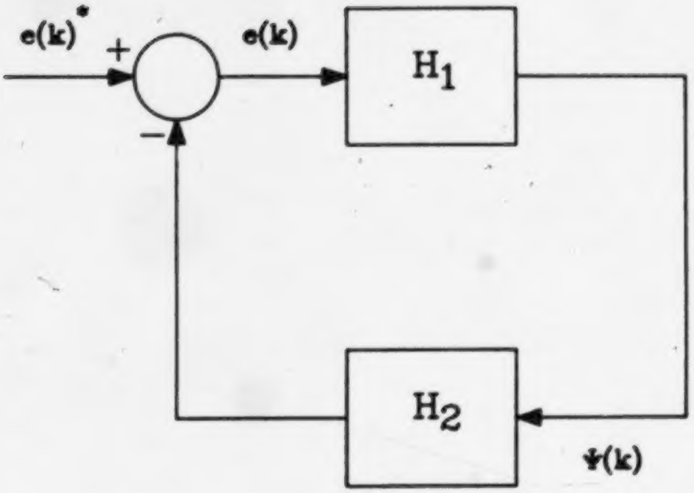
e(k)*
+
−
e(k)
H1
H2
Ψ(k)

Figure 4.16 Ortega *et al*'s error model schematic

n_r and μ

(i) $H_2[(\mu^{1/2}q)^{-1}]$ is stable and is strictly inside the cone$(\bar{\sigma}^{-1},\bar{\sigma}^{-1})$

(ii) $r(k) \epsilon L_\infty$ such that $e(k)^* \epsilon L_2$

then $\Psi(k)$, $e(k) \epsilon L_2$ and $\phi(k) \epsilon L_\infty$.

The center and radius of the cone in (i) are determined by the parameter $\bar{\sigma}$ which is greater than or equal to f, the adaptive gain in (4.54). If the Nyquist locus of $H_2(\mu^{1/2}q)$ lies strictly inside the circle in the right-half plane centered at $\bar{\sigma}^{-1}$ with radius $\bar{\sigma}^{-1}$ then the conic condition in (i) is satisfied (Safanov, 1980). This corresponds to a SPR type of condition. Note that for the matched case, H_2, which is equal to unity when $\tilde{G}=0$, lies strictly inside all circles with $f \le \bar{\sigma} < 2$.

4.4.2 Proposed Design Approach

The design approach proposed here consists of the following steps:

(i) derive H_2 for a given model-plant mismatch;

(ii) check if H_2 is stable;

(iii) plot the Nyquist locus of H_2 to ensure that it lies strictly in the right-half plane;

(iv) select values for f and μ which satisfy the conic restriction;

(v) select alternate design parameters (e.g. different model order or sampling rate) and iterate through steps (i) to (iv).

Derivation of H_2 in (4.60) requires a tuned parameter set θ^* as defined by the identity (4.59) and knowledge of the model uncertainty $\tilde{G}$ in (4.51). Using the true values for the plant parameters in the calculation of H_2 enables an exact Nyquist locus to be calculated. However if the true values are not known, which is usually the case, then estimates of these parameters along with error bands on the estimates may be used to calculate H_2. This would normally lead to a band of H_2 Nyquist loci calculated using the limiting values of the parameters, e.g. $\theta_i \pm \Delta_i$ where θ_i is the estimated value of the i-th plant parameter and Δ_i is the uncertainty. The stability analysis is unchanged except that the stability criterion requires that this Nyquist band must lie completely in the right-half plane.

The set of stabilizing parameters θ^* is not unique. Kosut and Johnson (1984) suggest that B/A in (4.51) should represent a *tuned* model that provides a good fit of the true plant at low frequencies, i.e. $\tilde{G}(0)=0$, $H_2(0)=1$. Alternatively, the fit may be made at some user-selected intermediate frequency. Once a set of θ^* values is obtained it can be improved by empirically perturbing each element of θ^* and selecting the set that gives the smallest magnitude for the H_2 locus in order that the stability conditions are not overly conservative. Once H_2 has been selected then steps (ii) through (v) are completed as illustrated by the example in the next section.

It is important to note that the stability analysis does not prove or claim that the parameter estimates $\hat{\theta}$ converge to the selected set of θ^* values, i.e. $\Psi(k)\epsilon L_2$ may be satisfied by orthogonality ($\hat{\theta}^t\phi \rightarrow \theta^{*t}\phi$). This is also true in the matched case (Goodwin *et al*, 1980) where $\hat{\theta}$ does not necessarily converge to the unique set of true process parameters.

To demonstrate the proposed design approach Rohrs' (1982) benchmark example from (4.38) is used with a first order model. To obtain tuned estimates of the coefficients of the A and B polynomials an open loop identification test was performed. For instance, at T_s=.04s, the discrete representation of the plant in (4.38) is given by

$$\frac{y(k)}{u(k)} = \frac{.361\times10^{-2}q^{-1}+.107\times10^{-1}q^{-2}+.194\times10^{-2}q^{-3}}{1-.205\times10^{1}q^{-1}+.135\times10^{1}q^{-2}-.289q^{-3}} \qquad (4.61)$$

Figure 4.17 shows the results of an identification run performed on the plant in (4.61). The CG/PAA was used for adjusting the two parameter estimates in the first order model.

$$\hat{y}(k)=\hat{b}_1u(k-1)+\hat{a}_1y(k-1) \qquad (4.62)$$

$$\hat{b}_1=0.076, \quad \hat{a}_1=0.962$$

$\tilde{G}$ may then be evaluated by equating the discrete plant in (4.61) to the process representation in (4.51) with A and B replaced by the tuned parameters in (4.62). For simplicity of analysis the desired closed loop polynomial has been selected as C_R=1. The polynomials S^* and R^* which satisfy identity (4.59) for the tuned model are

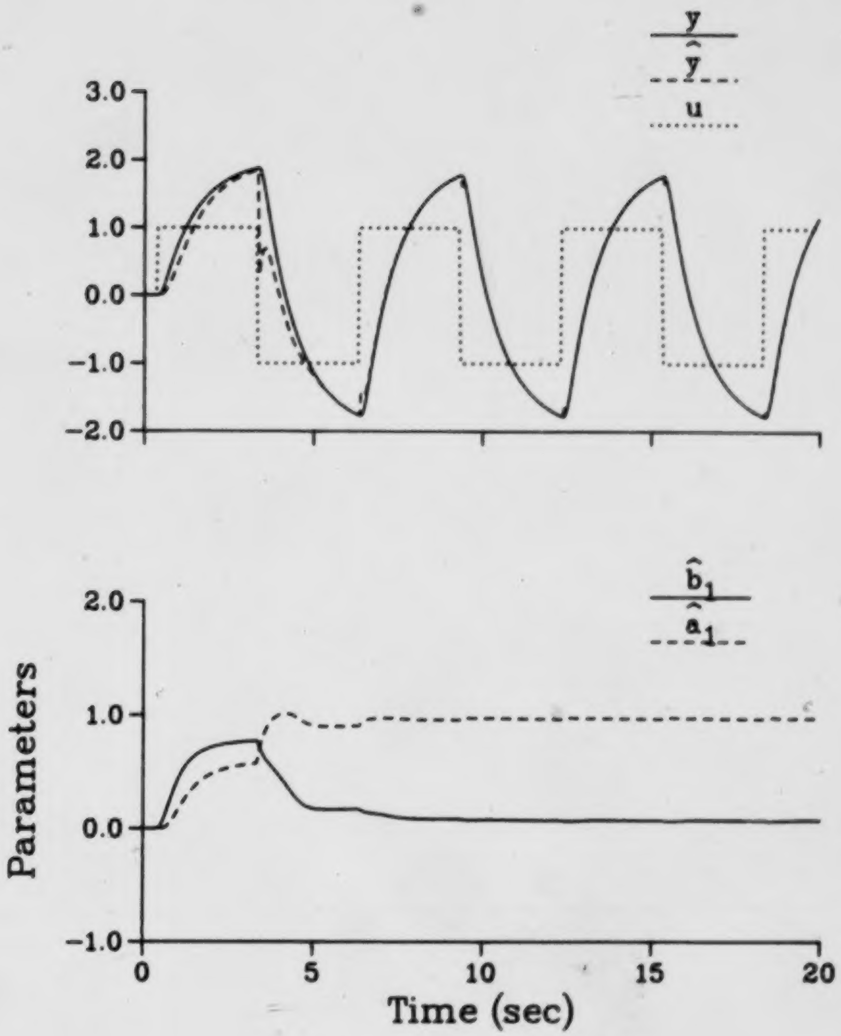

Figure 4.17 Convergence trajectories for first order model parameters

$S^*=\hat{b}_1,\ R^*=\hat{a}_1$

The result of substituting these values for $\tilde{G}$ and R^* into (4.60) is

$T=\hat{a}_1q^{-1}$

and an $H_2(q)$ of the form

$$H_2(q) = \frac{c_0q^4+c_1q^3+c_2q^2+c_3q}{d_0q^4+d_1q^3+d_2q^2+d_3q+d_4} \tag{4.63}$$

where the c_i's and d_i's are functions of the plant and tuned model parameter values. This procedure was repeated for several discrete versions of (4.38) over the range T_s=.04s to .4s.

The next step is to check if H_2 is stable. At T_s=.04s, it was found that $H_2(\mu^{1/2}q)$ is unstable at $\mu=1$ and hence is unstable for all $\mu\epsilon(0,1)$. For the other discrete plants (T_s=.08s to .4s) it was found that $H_2(\mu^{1/2}q)$ is stable for $\mu>.88$.

To illustrate the roles of adaptive gain, order of model-plant mismatch and sampling rate in adaptive system robustness, $\tilde{G}$ and then H_2 are evaluated for a number of different cases. The role of other design tools, such as the choice of the desired closed loop pole locations (C_R), could also be investigated in a similar manner.

4.4.3 Adaptive Gain Effects

As mentioned earlier, in the matched case (i.e. $H_2=1$), the gain f has an upper bound of 2. If the plant is represented by a 1st order model with T_s=0.32s, then the corresponding H_2 locus is as shown in Figure 4.18. This H_2 is strictly inside the cone($\bar{\sigma}^{-1}$,$\bar{\sigma}^{-1}$) with $f \leq \bar{\sigma} = 1$, as represented by the dotted line. Note that as the MPM increases the magnitude of the H_2 locus increases. The effect of increasing MPM is offset by decreasing the adaptive gain which increases the size of the dotted circle in the right-half plane. However, there is an upper limit on the amount of MPM that can be tolerated because if H_2 crosses into the left-half plane then there is no value for the adaptive gain which will guarantee stability.

Figure 4.18 indicates that an adaptive gain greater than unity should not be used in this case because the H_2 locus would no longer be strictly inside any dotted circle with a center and radius smaller than one. Therefore even with only a moderate amount of MPM the adaptive gain should be selected significantly less than the maximum gain of two allowed in the matched case.

4.4.4 Model Order Effects

In general it is obvious that a higher order model can capture or characterize more dynamics than a lower order model. Consequently $\tilde{G}$ for a higher order model should be smaller than $\tilde{G}$ for a lower order model. For example, the H_2 locus for a first order model with T_s=.24s is shown in Figure 4.19(a). A procedure identical to the approach described earlier was used to identify a second order model of the form

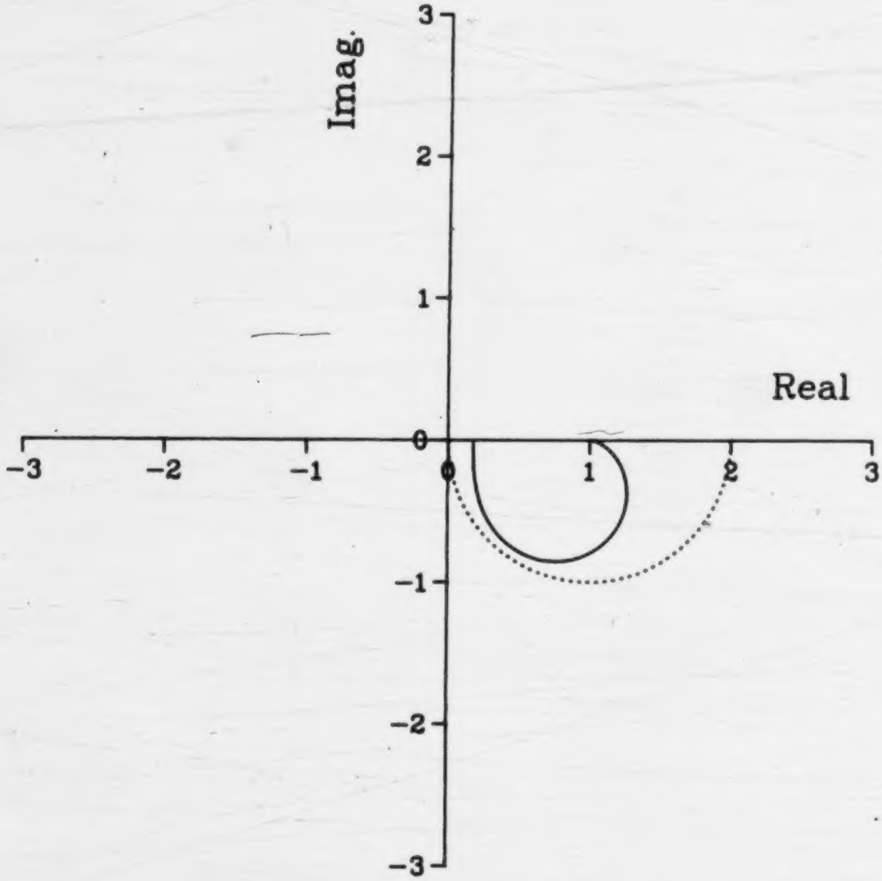

Figure 4.18 Nyquist locus of $H_2(\mu^{1/2}q)$ for a first order model (T_s=.32s, μ=0.9)

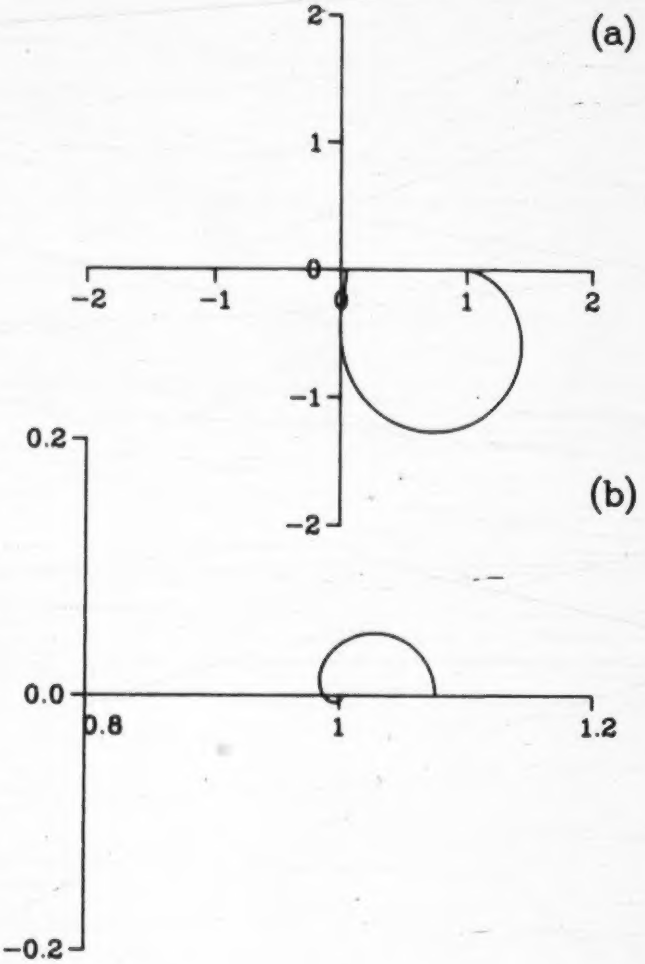

Figure 4.19 Nyquist locus of $H_2(\mu^{1/2}q)$ (T_s=0.24s, μ=0.9) for: (a) a first order model; and (b) a second order model

$$y(k)=\hat{b}_1u(k-1)+\hat{b}_2u(k-2)+\hat{a}_1y(k-1)+\hat{a}_2y(k-2) \qquad (4.64)$$

Converged parameter values from this test were

$\hat{b}_1=.219$, $\hat{b}_2=.177$, $\hat{a}_1=.859$, $\hat{a}_2=-.057$

The $H_2(\mu^{1/2}q)$ loci for this case with $T_s=.24s$ is plotted in Figure 4.19(b). By comparison with Figure 4.19(a), it is clear that the size of the H_2 locus is greatly reduced by selecting a second order model over a first order model. As demonstrated in subsection 4.4.3, the magnitude of H_2 and its closeness to unity are directly related to the amount of MPM, i.e. as $\tilde{G}\to 0$, $H_2\to 1$. Hence Figure 4.19 is an illustration of how model order selection can affect the robustness of an adaptive control system.

The choice of model made in this example represents one of many possibilities. Other candidates might be for instance a model with a delay different from the true plant.

4.4.5 Sampling Time Effects

The role of the sample period on the magnitude of H_2 is clearly evident from Figure 4.20 where the Nyquist loci of several $H_2(q)$ over the range $T_s=0.08s$ to $0.4s$ are plotted. As the sample time increases, the robustness increases in the sense that the magnitude of H_2 decreases and $H_2\to 1$. Hence the sampling period is an effective tool for offsetting the destabilizing effects of unmodeled dynamics. Choice of sampling time represents a degree of freedom available to the discrete system designer which is not available with continuous systems. However the robustness benefits of

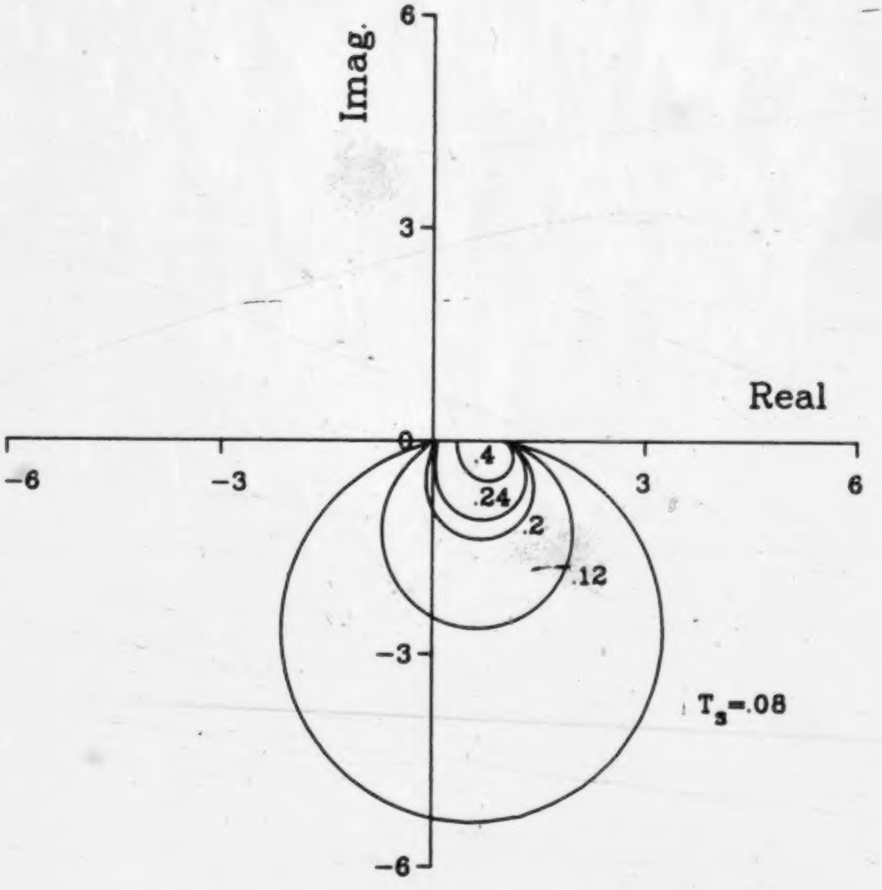

Figure 4.20 Nyquist loci of H_2 for first order models

increased sample period must be traded off against poorer performance due to reduced speed of parameter identification and quality of dynamic response.

From Figure 4.20, a suitable range for the sampling time is T_s=.3s to .4s with a first order model. This is well within common guidelines which suggest a sampling interval of .1 to .4 times the dominant time constant. This suggests that a choice of T_s towards the upper end of the range defined by typical guidelines is preferred if high frequency, unmodeled dynamics are present.

One last point worth mentioning here is that Ortega *et al*'s (1985) robust stability result does not assume explicitly that the process to be controlled is minimum phase. This is in contrast to the stable algorithms in Goodwin *et al* (1980) for example which have similar control laws and parameter adaptation schemes but require this assumption to prove stability. This raises an interesting question: is a reduced order control law based on the control scheme presented by Ortega *et al* (1985) capable of guaranteeing stability of a process which is nonminimum phase?

The conic sector results presented in this chapter provide some insight into this matter. From Figure 4.3 it was previously noted that the nonminimum phase characteristics do not disappear until T_s>0.2s. In Figure 4.20 it is observed that for T_s>0.2s the Nyquist loci of $H_2(q)$ lie strictly inside the right-half plane. Therefore there exist values for f and μ which would guarantee stability with the reduced order model. However for the discrete plants with unstable zeros ($T_s<0.2s$) the respective H_2 loci cross into the left-half plane and therefore stability cannot be guaranteed for these plants. Therefore for this example the adaptive control system of Ortega *et al* (1985) does not guarantee stability of the nonminimum phase plants with a reduced order model.

4.5 Conclusions

1. The nonminimum phase zero which results from fast sampling of Rohrs' example can cause stability problems even when no unmodeled dynamics are present. It is emphasized in this chapter that violation of the stable-inverse plant assumption plays as important a role as does the presence of unmodeled dynamics in causing instability.

2. It is demonstrated that the adaptive control algorithms studied by Rohrs have some robustness properties in the presence of unmodeled dynamics when the discrete plant representation is minimum phase and the sampling rate is selected properly.

3. The proposed design approach for discrete adaptive control systems with model-plant mismatch provides a quantitative measure of the effect of design alternatives such as model order on stability.

4.6 References

Astrom, K.J., P. Hagander and J. Sternby (1984). Zeros of sampled systems. *Automatica*, 20, 31-38.

Astrom, K.J. and B. Wittenmark (1984). *Computer Controlled Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Cluett, W.R., J.M. Martin-Sanchez, S.L. Shah and D.G. Fisher (1986). Stable discrete-time adaptive control in the presence of unmodeled dynamics. *1986 American Control Conference*, Seattle, WA, 713-720.

Doyle,J.C. and G. Stein (1981). Multivariable feedback design: concepts for a classical/modern synthesis. *IEEE Trans. Aut. Contr.*, AC-26, 4-16.

Egardt, B. (1979). *Stability of Adaptive Controllers*. Springer-Verlag, Berlin.

Goodwin,G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Kosut,R.L. and C.R. Johnson,Jr. (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Kreisselmeier, G. and B.D.O. Anderson (1986). Robust model reference adaptive control. *IEEE Trans. Aut. Contr.*, AC-31, 127-133.

Lim, K.W. (1982). Robustness of self-tuning controllers. Dept. of Eng. Sci., Oxford University, OUEL Report No. 1422/82.

Martin-Sanchez, J.M., S.L. Shah and D.G. Fisher (1984). A stable adaptive predictive control system. *Int. J. Contr.*, 39, 215-234.

Ortega,R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Rohrs, C. (1982). Adaptive control in the presence of unmodeled dynamics. Dept. of Elec. Eng. Comput. Sci., MIT, LIDS Report No. TH-1254.

Rohrs, C.E., M. Athans, L. Valavani and G. Stein (1984). Some design guidelines for discrete-time adaptive controllers. *Automatica*, 20, 653-660.

Rohrs, C.E., L. Valavani, M. Athans and G. Stein (1985a). Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics. *IEEE Trans. Aut. Contr.*, AC-30, 881-889.

Rohrs, C.E., G. Stein and K.J. Astrom (1985b). A practical robustness theorem for adaptive control. *1985 American Control Conference*, Boston, MA, 979-983.

Safanov,M.G. (1980). *Stability and Robustness of Multivariable Feedback Systems*, MIT Press, Cambridge, MA.

5. Stability Analysis of Discrete-Time Adaptive Control Systems based on Input-Output Theory

5.1 Introduction

Certainly one of the earliest attempts to analyze the robustness of discrete-time adaptive control systems was performed by Lim (1982). More specifically, Lim's work focused on the robustness of the Clarke and Gawthrop (1975) version of the self-tuning controller in the presence of unmodeled plant dynamics. Stability conditions were derived for the adaptive control system in terms of the design parameters available in the STC algorithm.

The stability approach taken by Lim (1982) closely followed the analysis performed by Gawthrop (1980) on the stability and convergence of a self-tuning controller when the process order is exactly equal to that of the assumed model (i.e. no unmodeled dynamics). Gawthrop (1980) cast the adaptive control system into an error feedback system. The stability of this feedback system was analyzed using input-output stability techniques. Gawthrop derived input-output properties for a particular relation in the feedback system based on a least squares type of parameter estimator. It is this same approach that was used by Lim (1982) to analyze the robustness of the STC algorithm.

The work of Gawthrop (1980) and Lim (1982) contains one major flaw. Both of their stability results use an *a priori* plant signal boundedness assumption. Ortega *et al* (1985) have recently extended the robustness analysis of Lim (1982) to include normalized signals in the parameter adaptation algorithm. This modification removes the defect in the earlier work of Gawthrop and Lim.

The results of Lim (1982) and Ortega *et al* (1985) represent some of the earliest work on the robustness of discrete-time adaptive controllers to the modeling assumption. The objectives of this chapter are 1) to present a critical review of this early work, 2) to illustrate and correct some of the errors in this work, and 3) to extend the results in such a way as to unify the work of Lim and Gawthrop, and Ortega *et al*.

5.2 Nonzero Initial Conditions

Gawthrop (1980), Lim (1982) and Ortega *et al* (1985) in their stability analyses derive sector properties for a relation in an error feedback system based on a least squares parameter adaptation algorithm. To illustrate these results the input-output properties of a relation based on a particular interlaced least squares algorithm are derived in Lemma 5.2.

Define

$$\Psi(k) = (\hat{\theta}(k-d)-\theta)^t\phi(k-d) = \tilde{\theta}(k-d)^t\phi(k-d) \qquad (5.1)$$

where $\tilde{\theta}$ is the difference between the parameter estimates used in the control law and an arbitrary vector θ and ϕ is a regressor vector containing past plant input and output signals. (Ortega *et al* (1985) in their error system analysis have defined θ as a vector of stabilizing parameters. However, the input-output properties of the above-mentioned relation may be derived independent of the properties of this vector θ.)

Consider the following update mechanism for the parameter estimates

$$\hat{\theta}(k) = \hat{\theta}(k-d) + F(k)\phi(k-d)e(k) \tag{5.2}$$

where e(k) is the tracking error. For a least squares type algorithm F(k) is a time-varying matrix adjusted according to

$$F(k)^{-1} = \beta F(k-d)^{-1} + \phi(k-d)\phi(k-d)^t \tag{5.3}$$

where β ($0<\beta\leq 1$) is the forgetting factor. A useful relationship between F(k) and F(k-d), to be used in Lemma 5.2, is presented in Lemma 5.1.

Lemma 5.1: If $F(k)^{-1}$ is updated according to (5.3) then

$$\phi(k-d)^t F(k)\phi(k-d) = \frac{\phi(k-d)^t F(k-d)\phi(k-d)}{\beta+\phi(k-d)^t F(k-d)\phi(k-d)} \tag{5.4}$$

Proof: From the Matrix Inversion Lemma, F(k) in (5.3) may be expressed as

$$F(k) = \beta^{-1}\left[F(k-d) - \frac{F(k-d)\phi(k-d)\phi(k-d)^t F(k-d)}{\beta+\phi(k-d)^t F(k-d)\phi(k-d)}\right] \tag{5.5}$$

Postmultiplying by $\phi(k-d)$

$$F(k)\phi(k-d) = \beta^{-1}\left[1 - \frac{\phi(k-d)^t F(k-d)\phi(k-d)}{\beta+\phi(k-d)^t F(k-d)\phi(k-d)}\right]F(k-d)\phi(k-d) \tag{5.6}$$

and the proof is completed by premultiplying by $\phi(k-d)^t$.

Lemma 5.2: The relation $H_1 : e(k) \rightarrow \Psi(k)$ defined by (5.2) and (5.3) is *weakly* $(1,1,\sigma_{LS})$ *dissipative* for all σ_{LS} satisfying

$$\sigma_{LS} > \frac{\phi(k-d)^t F(k-d)\phi(k-d)}{\beta+\phi(k-d)^t F(k-d)\phi(k-d)} \tag{5.7}$$

Proof: Consider the quadratic function

$$V(k)=\tilde{\theta}(k)^t F(k)^{-1}\tilde{\theta}(k) \tag{5.8}$$

$$=[\tilde{\theta}(k-d)+F(k)\phi(k-d)e(k)]^t \cdot F(k)^{-1} \cdot [\tilde{\theta}(k-d)+F(k)\phi(k-d)e(k)] \tag{5.9}$$

Replacing $F(k)^{-1}$ by (5.3) gives

$$V(k)=\beta V(k-d)+\Psi(k)^2+2\Psi(k)e(k)+\phi(k-d)^t F(k)\phi(k-d)e(k)^2 \tag{5.10}$$

Summing from 0 to N

$$\sum_{k=0}^{N}[V(k)-V(k-d)+(1-\beta)V(k-d)] = \sum_{k=0}^{N}[\Psi(k)^2+2\Psi(k)e(k)+\phi(k-d)^t F(k)\phi(k-d)e(k)^2] \tag{5.11}$$

Since $(1-\beta)\geq 0$ by definition and using the result of Lemma 5.1, it follows that

$$\sum_{k=0}^{N}[\Psi(k)^2+2\Psi(k)e(k)+\sigma_{LS}e(k)^2] \geq -\sum_{k=-1}^{-d} V(k) \tag{5.12}$$

which completes the proof of this lemma.

Both Gawthrop (1980) and Ortega *et al* (1985) in their analysis of similar parameter adaptation schemes concluded that $H_1 : e \rightarrow \Psi$ is outside the cone$(-1,(1-\sigma_{LS})^{1/2})$. However from the definition of a sector given by Safanov (1980), H_1 is only outside the cone$(-1,(1-\sigma_{LS})^{1/2})$ if the righthand side of (5.12) is greater than or equal to zero. This would require that the initial parameter errors $(\tilde{\theta})$ be equal to zero. As stated by Lim (1982), it is possible, in principle, to select initial conditions such that $V(k)=0 \; \forall \; k<0$. This is true if the vector θ in (5.1) is in fact arbitrary which is the case if the input-output properties are derived independent of the stability analysis. However, both Lim (1982) and Ortega *et al* (1985) require in their stability theorems that θ satisfy some specific condition and therefore cannot be arbitrary. As a result, the initial parameters are, in general, nonzero and hence (5.12) does not satisfy Safanov's (1980) exterior conic condition. However as proven in Lemma 5.2, the relation H_1 does satisfy the definition of a weakly dissipative system as presented by Hill and Moylan (1983).

The obvious question which arises is: does this flaw in the stability analyses of Lim, Gawthrop and Ortega alter in any way their final results? The following theorem accomodates the nonzero initial conditions and demonstrates that the final results may remain unchanged.

Theorem 5.1: Consider the following feedback system

$$\begin{aligned} e(k) &= u(k) - x(k) \\ y(k) &= H_1 e(k) \qquad (5.13) \\ x(k) &= H_2 y(k) \end{aligned}$$

with $H_1, H_2 : L_{2e} \rightarrow L_{2e}$ and $x(k), y(k), e(k) \in L_{2e}$ and $u(k) \in L_2$.

If

(a) $H_1 : e(k) \rightarrow y(k)$ satisfies

$$\sum_{k=0}^{N} [y(k)^2 + \alpha e(k) y(k) + \beta e(k)^2] \geq -\gamma \qquad (5.14)$$

and (b) $H_2 : y(k) \rightarrow x(k)$ satisfies

$$\sum_{k=0}^{N} [\beta x(k)^2 - \alpha x(k) y(k) + y(k)^2] \le -\eta \| (x(k), y(k)) \|_N^2 \qquad (5.15)$$

for some $\alpha, \beta, \gamma, \eta \in R$ and $\gamma, \eta > 0$,

then the closed loop signals $x(k), y(k) \in L_2$.

Proof: From (5.14) and using $e(k)=u(k)-x(k)$

$$\sum_{k=0}^{N} [\beta x(k)^2 - \alpha x(k) y(k) + y(k)^2] + \sum_{k=0}^{N} [\alpha u(k) y(k) - 2\beta u(k) x(k) + \beta u(k)^2] \ge -\gamma \qquad (5.16)$$

Combining (5.15) and (5.16)

$$-\eta \| (x(k), y(k)) \|_N^2 + \sum_{k=0}^{N} [\alpha u(k) y(k) - 2\beta u(k) x(k) + \beta u(k)^2] \ge -\gamma \qquad (5.17)$$

Using the Schwarz Inequality

$$\eta \| (x(k), y(k)) \|_N^2 - |\alpha| \cdot \| u(k) \| \cdot \| y(k) \|_N - 2|\beta| \cdot \| u(k) \| \cdot \| x(k) \|_N \le \gamma + |\beta| \cdot \| u(k) \|^2 \qquad (5.18)$$

Assume $\| (x(k), y(k)) \|_N^2 \rightarrow \infty$ as $N \rightarrow \infty$. Therefore, from (5.18)

$$\eta \le 0$$

This is a contradiction. Therefore $\| (x(k), y(k)) \|_N^2$ is bounded (i.e. $x(k), y(k) \in L_2$) which completes the proof of this theorem.

Theorem 5.1 is an L_2 stability result which accomodates the input-output properties of the parameter adaptation algorithm. For instance, note that $H_1 : e(k) \rightarrow \Psi(k)$ defined by (5.1-5.3) satisfies (a) of Theorem 5.1 with $\alpha = 2$ and $\beta = \sigma_{LS}$. According to Safanov (1980), H_2^{-1} is strictly inside the cone (C_1, R_1) if condition (b) is satisfied, where

$C_1 = 1$

$R_1 = (1-\sigma_{LS})^{1/2}$

From the sector properties stated by Safanov (1980), the condition on H_2^{-1} is equivalent to H_2 being strictly inside the cone(C_2,R_2) where

$C_2 = \sigma_{LS}^{-1}$

$R_2 = \sigma_{LS}^{-1}(1-\sigma_{LS})^{1/2}$

The error feedback system developed by Ortega *et al* (1985) is in a form identical to (5.13). The condition on H_2 in Theorem 5.1 is identical to the condition on the corresponding H_2 in the L_2 stability result of Ortega *et al* (1985). Therefore, it may be concluded that, despite the oversight with respect to the nonzero initial conditions of the adaptation algorithm, the end result has remained unchanged.

5.3 The Vanishing Radius Problem

Gawthrop (1980) has demonstrated that for a least squares type of parameter adaptation algorithm, σ_{LS} of (5.7) is strictly less than unity only if it is assumed *a priori* that the input-output vector ϕ is bounded. This result may be clearly seen from (5.7) where F is the positive definite matrix defined in (5.5). The robustness analysis then follows from Theorem 5.1 where H_2 is required to be strictly inside the cone(C_2,R_2) to ensure that the system is stable. There is a fundamental flaw in this argument. Note that plant signal boundedness needs to be assumed *a priori* in order to apply the stability theory to prove signal boundedness. If this *a priori* assumption is not made, then unity is the smallest value that may be chosen for σ_{LS}. Therefore, from Theorem 5.1, stability is ensured if H_2 is strictly inside the cone(1,0).

In both Lim's (1982) and Ortega *et al*'s (1985) robustness analyses, the relation H_2 contains information on the model-process mismatch. For the case when no mismatch is present (i.e. the model order is equivalent to the process order) the relation H_2 reduces to a simple scalar. For example, $H_2=1$ when no unmodeled dynamics are present in Ortega *et al* (1985). However, $H_2=1$ is not strictly inside the cone(1,0) and therefore stability cannot be concluded even in the matched case.

This behaviour of the allowable cone for H_2 is referred to as the vanishing radius problem, i.e. the conic region, without an *a priori* signal boundedness assumption, is reduced to a cone centered at unity with a zero radius. This problem arises in most of the early literature dealing with the robustness of discrete-time adaptive controllers based on input-output stability theory (Lim, 1982; Gawthrop and Lim, 1982; Ortega and Landau, 1983a,b). In the more recent work of Ortega *et al* (1985) the vanishing radius problem has been successfully addressed. Certainly one means of preventing the radius of the allowable cone for H_2 from disappearing would be to use a vector ϕ in the adaptation scheme which remains bounded. Ortega *et al* (1985) have considered normalizing the vector ϕ and using this normalized vector in the adaptation update algorithm. The normalization factor selected by Ortega was first introduced by Egardt (1979) and is of the form

$$\rho(k) = \mu\rho(k-1) + \max(|\phi(k-d)|^2,\bar{\rho}) \qquad (5.19)$$

$$\bar{\rho}>0, \ \mu\epsilon(0,1)$$

and the normalized input-output vector has been defined as

$$\phi^n(k-d) = \rho(k)^{-1/2}\cdot\phi(k-d) \qquad (5.20)$$

It may easily be shown that $|\phi^n(k-d)|^2<1 \ \forall k$, i.e.

$$|\phi^n(k-d)|^2 = \frac{|\phi(k-d)|^2}{\rho(k)}$$

$$= \frac{|\phi(k-d)|^2}{\mu\rho(k-1)+\max(|\phi(k-d)|^2,\bar{\rho})}$$

$$< 1$$

Ortega *et al* (1985) have also defined a normalized tracking error and a normalized $\Psi(k)$, i.e.

$$e^n(k) = \rho(k)^{-1/2}\cdot e(k) \qquad (5.21)$$

$$\Psi^n(k) = \rho(k)^{-1/2}\cdot\Psi(k) \qquad (5.22)$$

The parameter adaptation scheme in (5.2) and (5.3) is now expressed in terms of the normalized variables, i.e.

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + F(k)\phi^n(k-d)e^n(k) \qquad (5.23)$$

$$F(k)^{-1} = \beta F(k-d)^{-1} + \phi^n(k-d)\phi^n(k-d)^t \qquad (5.24)$$

Consider the above algorithm with $\beta=1$. From (5.5) it follows that

$$\lambda_{max}F(k) \le \lambda_{max}F(k-d) \qquad (5.25)$$

Therefore if $\lambda_1=\lambda_{max}F(0)$, where $F(0)$ is the initial value for the time-varying matrix F(k), then

$$\lambda_{max}F(k) \leq \lambda_1 \quad \forall k \tag{5.26}$$

Lemma 5.3: The relation $H_1^n : e^n(k) \rightarrow \Psi^n(k)$ defined by (5.23) and (5.24) with $\beta=1$ is weakly $(1,1,\sigma_{LS})$ dissipative for all σ_{LS} satisfying

$$\sigma_{LS} > \lambda_1/(1+\lambda_1) \tag{5.27}$$

Proof: Combining the result of Lemma 5.1 with the $F(k)^{-1}$ update in (5.24) gives

$$\phi^n(k-d)^t F(k) \phi^n(k-d) = \frac{\phi^n(k-d)^t F(k-d) \phi^n(k-d)}{1+\phi^n(k-d)^t F(k-d) \phi^n(k-d)} \tag{5.28}$$

Using (5.26) and the fact that $|\phi^n(k-d)|^2 < 1$, it may be concluded

$$\frac{\lambda_1}{1+\lambda_1} \geq \frac{\phi^n(k-d)^t F(k-d) \phi^n(k-d)}{1+\phi^n(k-d)^t F(k-d) \phi^n(k-d)} \tag{5.29}$$

The rest of the proof follows the same as in Lemma 5.2.

The value for λ_1 is selected by the user. As long as λ_1 is finite, σ_{LS} may be selected strictly less than unity. Therefore the vanishing radius problem has been avoided.

5.4 Choice of the Normalization Factor

Although the use of the normalized variables in the parameter adaptation algorithm has solved the vanishing radius problem, it has created another problem. Theorem 5.1, under the normalized system, only guarantees L_2 stability of the normalized variables. Does the stability of the normalized signals imply the stability of the unnormalized signals (which is what is of interest)? This problem was solved by Ortega *et al* (1985) for their particular choice of normalization factor (5.19) using the Bellman-Gronwall Lemma. This problem was also solved by Goodwin *et al* (1980) for a different choice of normalization factor using their Key Technical Lemma (KTL).

The result presented in Theorem 5.1 and the KTL are combined in Corollary 5.1 to provide an alternate proof of stability for the adaptive controller presented by Goodwin *et al* (1980). In doing so, it is demonstrated that one interpretation of the role of the KTL is to prove stability of the unnormalized signals from the stability of the normalized signals.

Consider that the plant may be written as

$$y(k) = \theta^t\phi(k-d) \tag{5.30}$$

where θ is a vector of the unknown process parameters and ϕ is a vector containing past values of the plant input, u, and output, y. For projection algorithm I of Goodwin *et al* (1980), the predictive control law is used where u(k) satisfies

$$r(k+d) = \hat{\theta}(k)^t\phi(k) \tag{5.31}$$

Consider the following d interlaced version of the parameter update scheme used by Goodwin *et al* (1980)

$$\hat{\theta}(k) = \hat{\theta}(k-d) + \frac{a(k)\phi(k-d)e(k)}{1+\phi(k-d)^t\phi(k-d)} \tag{5.32}$$

where $a(k)\epsilon(0,2)$ and e(k) is the tracking error defined as

$$e(k) = y(k)-r(k) \tag{5.33}$$

Define the quadratic function

$$V(k) = \tilde{\theta}(k)^t\tilde{\theta}(k)/2 \tag{5.34}$$

where $\tilde{\theta}$ is the parameter error vector, i.e. $\tilde{\theta}(k)=\hat{\theta}(k)-\theta$. Let the normalization factor be selected as

$$\rho(k) = a(k)^{-1}(1+\phi(k-d)^t\phi(k-d)) \tag{5.35}$$

Substituting (5.32) into (5.34) gives

$$V(k)-V(k-d) = \rho(k)^{-1}\Psi(k)e(k) + \rho(k)^{-2}\phi(k-d)^t\phi(k-d)e(k)^2/2 \tag{5.36}$$

where $\Psi(k)$ has been defined previously in (5.1). Define $\sigma(k)$ as

$$\sigma(k) = \rho(k)^{-1}\phi(k-d)^t\phi(k-d) \tag{5.37}$$

and let the normalized signals $e^n(k)$ and $\Psi^n(k)$ be defined as in (5.21), (5.22) with $\rho(k)$ given by (5.35). Therefore (5.36) may be rewritten as

$$V(k)-V(k-d) = \Psi^n(k)e^n(k) + \sigma(k)e^n(k)^2/2 \tag{5.38}$$

Summing (5.38) from 0 to N gives

$$\sum_{k=0}^{N}[\Psi^n(k)e^n(k) + \sigma(k)e^n(k)^2/2] \geq -\sum_{k=-1}^{-d} V(k) \tag{5.39}$$

A corollary of Theorem 5.1 is presented which will be useful for analyzing projection algorithm I of Goodwin *et al* (1980).

Corollary 5.1: Consider the feedback system of (5.13). If

(a) H_1:e(k)→y(k) satisfies

$$\sum_{k=0}^{N}[e(k)y(k)+\bar{\sigma}e(k)^2/2] \geq -\gamma \qquad (5.40)$$

and (b) H_2:y(k)→x(k) satisfies

$$\sum_{k=0}^{N}[\bar{\sigma}x(k)^2/2-x(k)y(k)] \leq -\eta||(x(k),y(k))||_N^2 \qquad (5.41)$$

for some $\bar{\sigma}$, γ, $\eta > 0$,

then the closed loop signals $x(k),y(k)\epsilon L_2$.

Proof: The proof follows the approach taken for Theorem 5.1. From (5.40) and using e(k)=u(k)-x(k)

$$\sum_{k=0}^{N}[\bar{\sigma}x(k)^2/2-x(k)y(k)] + \sum_{k=0}^{N}[u(k)y(k)+\bar{\sigma}u(k)^2/2-\bar{\sigma}u(k)x(k)] \geq -\gamma$$

and combining (5.41) gives

$$-\eta||(x(k),y(k))||_N^2 + \sum_{k=0}^{N}[u(k)y(k)+\bar{\sigma}u(k)^2/2-\bar{\sigma}u(k)x(k)] \geq -\gamma$$

Using the Schwarz Inequality

$$\eta||(x(k),y(k))||_N^2-||u(k)||\cdot||y(k)||_N -\bar{\sigma}||u(k)||\cdot||x(k)||_N \leq \gamma+\bar{\sigma}||u(k)||^2/2$$

Assuming $||(x(k),y(k))||_N^2\rightarrow\infty$ as $N\rightarrow\infty$ produces a contradiction (i.e. $\eta\leq 0$) and therefore $x(k),y(k)\epsilon L_2$ which completes the proof of this corollary.

If H_1^n:$e^n(k)\rightarrow\Psi^n(k)$, then from (5.39), H_1^n satisfies property (a) of Corollary 5.1 with $\bar{\sigma}>\sigma(k)$ $\forall k$. From their definitions, $\Psi(k)$ and e(k) are related as follows

$$\begin{aligned} e(k) &= y(k)-r(k) \\ &= \theta^t\phi(k-d) - \hat{\theta}(k-d)^t\phi(k-d) \quad \text{(using (5.30) and (5.31))} \\ &= -\tilde{\theta}(k-d)^t\phi(k-d) \\ &= -\Psi(k) \end{aligned} \tag{5.42}$$

Therefore the control system presented in (5.30-5.33) may be cast into the feedback system of (5.13), i.e.

$$\begin{aligned} \Psi^n(k) &= H_1{}^n e^n(k) \\ e^n(k) &= -H_2{}^n\Psi^n(k) \end{aligned} \tag{5.43}$$

with $H_1{}^n$ satisfying (5.39) and $H_2=H_2{}^n=1$. L_2 stability of $e^n(k)$ and $\Psi^n(k)$ is guaranteed if $H_2{}^n$ satisfies (b) of Corollary 5.1. Condition (b) may be rearranged as

$$\sum_{k=0}^{N}[x(k)^2-(2/\bar{\sigma})x(k)y(k)] \le -\eta'\,||(x(k),y(k))||_N{}^2 \tag{5.44}$$

where $\bar{\sigma}>0$ and $\eta'=2\eta/\bar{\sigma}$. According to Safanov (1980) $H_2:y(k)\to x(k)$ is strictly inside the cone$(\bar{\sigma}^{-1},\bar{\sigma}^{-1})$ if condition (5.44) is satisfied. From (5.35) and (5.37)

$$\begin{aligned} \sigma(k) &= \frac{a(k)\phi(k-d)^t\phi(k-d)}{1+\phi(k-d)^t\phi(k-d)} \\ &< a(k) < 2 \quad \text{(since } a(k)\epsilon(0,2)\text{)} \end{aligned} \tag{5.45}$$

Therefore it is possible to select a $\bar{\sigma}$ such that $\sigma(k)<\bar{\sigma}<2$ and then L_2 stability of (5.43) follows because $H_2{}^n=1$ is strictly inside the cone$(\bar{\sigma}^{-1},\bar{\sigma}^{-1})$. L_2 stability of the normalized signals implies that

$$e^n(k)^2 = \frac{e(k)^2}{\rho(k)} = \frac{a(k)e(k)^2}{1+\phi(k-d)^t\phi(k-d)} \to 0 \text{ as } k\to\infty \tag{5.46}$$

As done by Goodwin *et al* (1980), the KTL may be used in conjunction with (5.46) and a stable-inverse plant assumption to prove that $\{||\phi(k)||\}$ is bounded and $e(k)\epsilon L_2$ (i.e. stability of the unnormalized error signal).

Note that in the above analysis of Goodwin *et al*'s (1980) projection algorithm, the model order selected for the control law in (5.31) (i.e. the dimension of $\hat{\theta}$ and ϕ) is equal to the true process order. If any mismatch were introduced into the control system formulation presented in (5.30-5.33), it would be discovered that e(k) would no longer be equal to $-\Psi(k)$ as in (5.42) (i.e. $H_2{}^n \neq 1$).

Ortega *et al* (1985) have extended the feedback system of (5.43) to include process-model order mismatch and bounded disturbances to the plant output. The assumed plant to be controlled is described by

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + \xi(k) \qquad (5.47)$$

where A and B are polynomials in the backward shift operator, d is the delay and $\xi(k)$ is a bounded disturbance sequence. The tracking error is defined as

$$e(k) = C_R y(k) - r(k) \qquad (5.48)$$

where C_R describes the desired closed loop poles. The regulator structure is based on a predictive control law and is given by

$$r(k+d) = \hat{S}(k)u(k) + \hat{R}(k)y(k) \qquad (5.49)$$

where $\hat{S}$ and $\hat{R}$ are polynomials in q^{-1} of degrees n_s and n_r with time-varying coefficients. This control law may be written in vector notation as

$$r(k+d) = \hat{\theta}(k)^t\phi(k) \qquad (5.50)$$

where

$$\phi(k)^t = [u(k)\cdots u(k-n_s), y(k)\cdots y(k-n_r)]$$

(Note that the control law in (5.50) is very similar to the one in (5.31). However the dimension of ϕ in (5.50) may be less than that of the I/O vector in the actual plant description.) Ortega *et al* (1985) then introduce a stabilizability assumption. A vector of parameters θ^* is defined as

$$\theta^{*t} = [s_0^* \; s_1^* \cdots r_0^* \; r_1^* \cdots] \tag{5.51}$$

where s_i^* and r_i^* are the constant coefficients of the polynomials S^* and R^* in q^{-1} of degrees n_s and n_r, respectively. The polynomial C is defined as

$$C = AS^* + q^{-d}R^*B \tag{5.52}$$

Ortega *et al* (1985) then assume that there exists a nonempty set θ_{LS} such that

$$\theta_{LS} = \{\theta^*: C(q) \neq 0, |q| > \mu^{1/2}\} \neq \emptyset \tag{5.53}$$

where $\mu \epsilon (0,1)$. This set θ_{LS} defines the fixed parameter controllers which produce closed loop poles within a disk of radius $\mu^{1/2}$. From (5.47)-(5.53), Ortega *et al* (1985) derive an error feedback system.

$$\begin{aligned} e(k) &= -H_2\Psi(k) + e(k)^* \\ \Psi(k) &= H_1 e(k) \end{aligned} \tag{5.54}$$

where

$$e(k)^* = (H_2 - 1)r(k) + C_R C^{-1} S^* \xi(k)$$

$$H_2 = C_R C^{-1} B$$

(Note that in the matched case where there is no model-process order mismatch, there exists S^* and R^* such that $C = C_R B$ and hence $H_2 = 1$.)

Ortega *et al* (1985) introduce normalized variables at this point (see (5.19)-(5.22)) and proceed to derive the I/O properties of $H_1{}^n$ for two parameter adaptation algorithms; 1) a constant gain (CG) algorithm, and 2) a regularized least squares (RLS) algorithm. The CG scheme is of the form

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + f\phi^n(k-d)e^n(k) \tag{5.55}$$

$$f \epsilon R, \; f>0$$

If (5.20) and (5.21) are substituted into (5.55), then

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + \frac{f\phi(k-d)e(k)}{\mu\rho(k-1)+\max(|\phi(k-d)|^2,\bar{\rho})} \tag{5.56}$$

(5.56) differs from the projection algorithm used by Goodwin *et al* (1980) where

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + \frac{a(k)\phi(k-d)e(k)}{1+|\phi(k-d)|^2} \tag{5.57}$$

The RLS scheme is given by

$$\tilde{\theta}(k) = \tilde{\theta}(k-d) + F(k)\phi^n(k-d)e^n(k) \tag{5.58}$$

$$F(k)=(1-\lambda_0/\lambda_1)[F(k-d)-\frac{F(k-d)\phi^n(k-d)\phi^n(k-d)^tF(k-d)}{\lambda+\phi^n(k-d)^tF(k-d)\phi^n(k-d)}]+\lambda_0 I \tag{5.59}$$

where $\lambda_0<\lambda_1$, λ are strictly positive scalars. The eigenvalues of $F(k)$ are contained in the interval $[\lambda_0,\lambda_1]$. The L_2 stability analysis of the RLS scheme may go along without the regularization (i.e. $\lambda_0=0$) which reduces the $F(k)$ update to

$$F(k)=F(k-d)-\frac{F(k-d)\phi^n(k-d)\phi^n(k-d)^tF(k-d)}{\lambda+\phi^n(k-d)^tF(k-d)\phi^n(k-d)} \tag{5.60}$$

It is proven by Ortega *et al* (1985) that for the CG algorithm defined by (5.55), $(H_1{}^n+\sigma_{CG}/2)$ is passive for all $\sigma_{CG}>f$. The proof follows exactly the same as the analysis required to derive (5.39) for the projection algorithm.

The I/O properties of $H_1{}^n$ for the RLS algorithm are rederived because there is an error in the proof presented by Ortega *et al* (1985).

Lemma 5.4: The relation $H_1{}^n:e^n(k)\to\Psi^n(k)$ defined by (5.58) and (5.60) is weakly $(1,\lambda,\lambda^2\cdot\sigma_{RLS})$ dissipative for σ_{RLS} satisfying

$$\sigma_{RLS} > \lambda_1/(\lambda+\lambda_1) \tag{5.61}$$

where $\lambda_1=\lambda_{max}F(0)$.

Proof: Consider the quadratic function

$$V(k) = \tilde{\theta}(k)^t F(k)^{-1}\tilde{\theta}(k) \tag{5.62}$$

From the Matrix Inversion Lemma, (5.60) may be reexpressed as

$$F(k)^{-1} = F(k-d)^{-1} + (1/\lambda)\cdot\phi^n(k-d)\phi^n(k-d)^t \tag{5.63}$$

Substituting (5.58) and (5.63) into (5.62) gives

$$\begin{aligned} V(k) = {} & V(k-d)+(1/\lambda)\cdot\Psi^n(k)^2+2\Psi^n(k)e^n(k) \\ & +\phi^n(k-d)^tF(k)\phi^n(k-d)e^n(k)^2 \end{aligned} \tag{5.64}$$

From (5.60) and following the proof of Lemma 5.1 it may be shown that

$$\phi^n(k-d)^tF(k)\phi^n(k-d) = \frac{\lambda\phi^n(k-d)^tF(k-d)\phi^n(k-d)}{\lambda+\phi^n(k-d)^tF(k-d)\phi^n(k-d)} \tag{5.65}$$

Then combining (5.65) with the fact that $|\phi^n(k-d)|^2<1$

$$\lambda\cdot\lambda_1/(\lambda+\lambda_1) > \phi^n(k-d)^tF(k)\phi^n(k-d) \tag{5.66}$$

Multiplying (5.64) through by λ and summing from 0 to N gives

$$\sum_{k=0}^{N}[\Psi^n(k)^2+2\lambda\Psi^n(k)e^n(k)+(\lambda^2\cdot\lambda_1/(\lambda+\lambda_1))\cdot e^n(k)^2]$$
$$\geq -\lambda\cdot\sum_{k=-1}^{-d} V(k) \qquad (5.67)$$

which completes the proof of this lemma.

L_2 stability of the normalized error system based on the CG algorithm is ensured if property (b) of Corollary 5.1 is satisfied with $\sigma_{CG}=\bar{\sigma}$, i.e. if $H_2{}^n$ of (5.54) is strictly inside the cone$(\sigma_{CG}{}^{-1},\sigma_{CG}{}^{-1})$. L_2 stability based on the simplified RLS algorithm (5.60) is ensured if property (b) of Theorem 5.1 is satisfied with $\alpha=2\lambda$ and $\beta=\lambda^2\cdot\lambda_1/(\lambda+\lambda_1)$, i.e. if $(H_2{}^n)^{-1}$ is strictly inside the cone(C_3,R_3) where

$$C_3 = \lambda$$
$$R_3 = \lambda(1-\sigma_{RLS})^{1/2}$$

or equivalently if $H_2{}^n$ is strictly inside the cone(C_4,R_4) where

$$C_4 = (\lambda\cdot\sigma_{RLS})^{-1}$$
$$R_4 = (\lambda\cdot\sigma_{RLS})^{-1}\cdot(1-\sigma_{RLS})^{1/2}$$

This last condition on $H_2{}^n$ differs from the result stated by Ortega *et al* (1985) where λ was omitted from C_4 and R_4. If $\lambda=1$ in (5.60) then the two results are the same.

The appearance of λ in the allowable cone for $H_2{}^n$ puts restrictions on the value selected for λ. For instance, in the matched case (5.42), $H_2=H_2{}^n=(H_2{}^n)^{-1}=1$. If $\lambda<1$, then C_3+R_3 must be greater than unity in order that $(H_2{}^n)^{-1}$ be strictly inside the cone(C_3,R_3). If $\sigma_{RLS}\simeq 0$ (i.e. $\lambda_1\simeq 0$), then λ must be greater than 0.5. If $\sigma_{RLS}\rightarrow 1$ (i.e. $\lambda_1>>1$), then there are no values of $\lambda<1$ such that the conicity condition on $(H_2^n)^{-1}$ is satisfied. A similar argument holds for $\lambda>1$.

As pointed out in section 5.3, the use of normalized variables in a least squares type of algorithm (e.g. RLS) prevents the vanishing radius problem (i.e. σ_{RLS} may be selected to be strictly less than unity). However the normalization factor selected by Ortega *et al* (1985) to avoid this problem (5.19) is not unique, i.e. the only requirement for the normalization is that the norm of ϕ^n (5.20) be bounded. For example, several variations of the normalization factor in (5.35) would serve this purpose. The question that arises is: why did Ortega *et al* (1985) select the normalization factor in (5.19)?

Some explanation for their choice of normalization is given by Ortega *et al* (1985). The factor in (5.19) was originally introduced by Egardt (1979). Ortega *et al* (1985) state that its importance with respect to robustness was established by Praly (1983) where it was proposed that the unmodeled dynamics be treated as an open loop disturbance. The normalization was required by Praly (1983) in order that boundedness of this disturbance was independent of an I/O boundedness assumption. What is not pointed out by Ortega *et al* (1985) is why other normalization factors such as (5.35) may not be as suitable for their analysis.

One important result in Ortega *et al* (1985) is the derivation of a conicity condition over H_2 which ensures that the conicity condition over H_2^n (e.g. cone(C_4,R_4)) is satisfied. This result is important because conditions on H_2 in (5.54) are more useful to the designer than conditions on H_2^n. It is hidden in the derivation of this result why normalization factors of the form in (5.35) are not as suitable for the conic sector type of stability analysis.

In general, $H_2 \neq H_2^n$ where $H_2 : x(k) \rightarrow y(k)$ and $H_2^n : x^n(k) \rightarrow y^n(k)$ with the normalized signals defined as in (5.21). However a special case is when $H_2 = c$, where c is a scalar. Then

$$y(k) = c \cdot x(k) \tag{5.68}$$

$$y^n(k) = H_2^n \cdot x^n(k) \tag{5.69}$$

Multiplying (5.68) through by $\rho(k)^{-1/2}$ gives

$$y^n(k) = c \cdot x^n(k) \tag{5.70}$$

From comparing (5.69) and (5.70) it may be concluded that $H_2^n = H_2 = c$. However, this is only the case when H_2 contains no dynamics. From the definition of H_2 in (5.54) it is clear that H_2 is, in general, a rational function in the backward shift operator q^{-1}. For example, let

$$H_2 = (1-q^{-1})^{-1} \tag{5.71}$$

Therefore

$$y(k) = x(k) + y(k-1) \tag{5.72}$$

Now assume that $H_2 = H_2^n$. Therefore

$$y^n(k) = x^n(k) + y^n(k-1) \tag{5.73}$$

Multiplying (5.73) through by $\rho(k)^{1/2}$ gives

$$y(k) = x(k) + \rho(k)^{1/2}\rho(k-1)^{-1/2}y(k-1) \tag{5.74}$$

(5.72) and (5.74) are equivalent if $\rho(k) = \rho(k-1)$. However, from inspection of (5.19) it may be seen that this is not true in general and therefore $H_2 \neq H_2^n$. The implications of this are that conditions on H_2^n from the L_2 stability analysis of the normalized error system cannot be interpreted directly as conditions on H_2 except for the special case mentioned above.

Ortega *et al* (1985) resolved this difficulty by deriving conic conditions on H_2 which ensure the conditions on H_2^n. This derivation is presented here to illustrate more clearly the role of the normalization factor.

Lemma 5.5(see also Lemma 5.1 of Ortega *et al* (1985)): Consider the relation $H:x(k) \rightarrow y(k)$. If $H[(\mu^{1/2}q)^{-1}]$ is inside the cone(C,R), then $H^n:x^n(k) \rightarrow y^n(k)$ with the normalized variables defined as in (5.19) and (5.21) is inside the same cone(C,R).

Proof: Define

$$w(k) = (y(k)-Cx(k))^2 - (Rx(k))^2 \tag{5.75}$$

$$w_1(k) = (y^n(k)-Cx^n(k))^2 - (Rx^n(k))^2 = \rho(k)^{-1}w(k) \tag{5.76}$$

Taking the sum from 0 to N of $w_1(k)$ gives

$$\sum_{k=0}^{N} w_1(k) = \sum_{k=0}^{N} \mu^k \rho(k)^{-1} \mu^{-k} w(k) \tag{5.77}$$

$$= \mu^{N+1}\rho(N+1)^{-1} \cdot \sum_{k=0}^{N} \mu^{-N} w(k) - \sum_{j=0}^{N} \left[\left(\sum_{k=0}^{j} \mu^{-k} w(k) \right) \cdot \left(\mu^{j+1}\rho(j+1)^{-1} - \mu^{j}\rho(j)^{-1} \right) \right] \tag{5.78}$$

(5.78) may be verified by straightforward expansion.

Multiplying (5.19) through by $\mu^{-(k+1)}$ gives

$$\mu^{-(k+1)}\rho(k) = \mu^{-k}\rho(k-1) + \mu^{-(k+1)} \cdot \max(|\phi(k-d)|^2, \bar{\rho}) \tag{5.79}$$

Therefore, $\mu^{-(k+1)}\rho(k)$ is nondecreasing (or $\mu^{k+1}\rho(k)^{-1}$ is nonincreasing) and hence in (5.78)

$$\mu^{j+1}\rho(j+1)^{-1} - \mu^{j}\rho(j)^{-1} < 0 \quad \forall j \tag{5.80}$$

Now express $H:x(k)\rightarrow y(k)$ as a general rational function in q^{-1}, i.e.

$$H(q^{-1}) = \frac{a_0+a_1q^{-1}+\cdots+a_nq^{-n}}{b_0+b_1q^{-1}+\cdots b_mq^{-m}} \tag{5.81}$$

Therefore $x(k)$ and $y(k)$ are related as follows

$$b_0y(k)+\cdots+b_my(k-m) = a_0x(k)+\cdots+a_nx(k-n) \tag{5.82}$$

Multiplying (5.82) through by $\mu^{-k/2}$ gives

$$\begin{aligned} &b_0\mu^{-k/2}y(k)+\cdots+b_m\mu^{-k/2}y(k-m) \\ &= a_0\mu^{-k/2}x(k)+\cdots+a_n\mu^{-k/2}x(k-n) \end{aligned} \tag{5.83}$$

Define

$$x'(k) = \mu^{-k/2}x(k) \tag{5.84}$$

$$y'(k) = \mu^{-k/2}y(k) \tag{5.85}$$

Substituting (5.84) and (5.85) into (5.83) allows us to write

$$\begin{aligned} &b_0y'(k)+b_1\mu^{-1/2}y'(k-1)+\cdots+b_m\mu^{-m/2}y'(k-m) \\ &= a_0x'(k)+\cdots+a_n\mu^{-n/2}x'(k-n) \end{aligned} \tag{5.86}$$

which may be expressed as

$$y'(k) = H[(\mu^{1/2}q)^{-1}]\cdot x'(k) \tag{5.87}$$

Therefore if $H[(\mu^{1/2}q)^{-1}]$ is inside the cone(C,R), then

$$\begin{aligned} &\sum_{k=0}^{N}[(y'(k)-Cx'(k))^2-(Rx'(k))^2] \\ &\quad = \sum_{k=0}^{N}[(\mu^{-k/2}y(k)-C\mu^{-k/2}x(k))^2-(R\mu^{-k/2}x(k))^2] \\ &\quad \le 0 \end{aligned} \tag{5.88}$$

which implies that

$$\sum_{k=0}^{N}\mu^{-k}w(k) \le 0 \tag{5.89}$$

From combining (5.89) and (5.80) with (5.78), it may be concluded that

$$\sum_{k=0}^{N} w_1(k) \le 0 \tag{5.90}$$

Therefore H^n is also inside the cone(C,R) which completes the proof of this lemma.

Specific properties of the selected normalization factor come into the proof of Lemma 5.5 only at (5.80). Let us select a normalization similar to (5.35) and see if it also satisfies a condition similar to (5.80). Let

$$\rho(k) = a(k)^{-1}(1+|\phi(k-d)|^2) \tag{5.91}$$

Since μ no longer appears in this factor the summation in (5.77) is written without μ as

$$\sum_{k=0}^{N} w_1(k) = \sum_{k=0}^{N} \rho(k)^{-1} w(k) \tag{5.92}$$

$$= \rho(N+1)^{-1} \sum_{k=0}^{N} w(k) \tag{5.93}$$

$$-\sum_{j=0}^{N} [(\sum_{k=0}^{j} w(k)) \cdot (\rho(j+1)^{-1} - \rho(j)^{-1})]$$

For the choice of $\rho(k)$ in (5.91).

$$\rho(j+1)^{-1} - \rho(j)^{-1} = a(j+1)(1+|\phi(j-d+1)|^2)^{-1}$$

$$-a(j)(1+|(j-d)|^2)^{-1} \tag{5.94}$$

The left hand side of (5.94) is less than or equal to zero if

$$\frac{a(j+1)}{a(j)} \le \frac{1+|\phi(j-d+1)|^2}{1+|\phi(j-d)|^2} \tag{5.95}$$

If (5.95) is satisfied and $\sum_{k=0}^{N} w(k) < 0$, then $\sum_{k=0}^{N} w_1(k) < 0$ and therefore H inside the cone(C,R) would imply H^n inside the same cone. (A similar analysis was performed by Ortega and Landau (1983) for a generalized least mean squares algorithm.) This approach does gives some guidelines for selecting the adaptation gain a(k) (e.g. (5.32)). However ensuring an inequality such as (5.95) may violate other restrictions on $a(k)$ (e.g. $a(k)\epsilon(0,2)$).

Some motivation for using a normalization factor of the form in (5.19) over (5.91) is now more clear. From the proof of Lemma 5.5, inequality (5.80) is guaranteed without any effect on the choice of design parameters such as adaptation gain, i.e. any choice of $\mu\epsilon(0,1)$ will lead to (5.80). On the other hand, ensuring (5.95) requires careful selection at each sampling instant of $a(k)$.

5.5 Extension to a L_∞ Result

All of the previous results presented in this chapter have been based on the L_2 stability result in Theorem 5.1. For the error system developed by Ortega *et al.* (1985) in (5.54), the L_2 stability theorem restricts $e(k)^*\epsilon L_2$ which in general requires r and ξ to eventually decay to zero. The more practical case is to consider bounded reference signals and disturbances (i.e. $r,\xi\epsilon L_\infty$).

Both Lim (1982) and Ortega *et al* (1985) considered exponentially weighted signals for their respective L_∞ analyses. For instance, define $\tilde{x}(k)$ as the exponentially weighted counterpart of $x(k)$ where

$$\tilde{x}(k) = \alpha^k \cdot x(k) \qquad \alpha>0 \tag{5.96}$$

(The importance of using weighted signals for the L_∞ case will be made more clear later in this section.) The next step performed by Lim (1982) and Ortega *et al* (1985) was to derive input-output properties of the parameter adaptation scheme based on these exponentially weighted signals. For example, let us examine the algorithm defined in (5.23) and (5.24).

Lemma 5.6: Consider the relation H_1^n: $e^n(k)\to\Psi^n(k)$ defined by (5.23) and (5.24). Define its exponentially weighted counterpart as $\bar{H}_1^n:\bar{e}^n(k)\to\bar{\Psi}^n(k)$. $\bar{H}_1^n$ is weakly $(1,1,\sigma_{LS})$ dissipative for σ_{LS} satisfying

$$\sigma_{LS} > \lambda_1/(\beta+\lambda_1) \tag{5.97}$$

where $\lambda_1=\lambda_{max}F(k)$ $\forall k$ if $(1-\beta a^{2d})\geq 0$.

Proof: Consider the quadratic function

$$V(k) = \tilde{\theta}(k)^t F(k)^{-1}\tilde{\theta}(k) \tag{5.98}$$

Using (5.10) in conjunction with the normalized signals gives

$$\begin{aligned} V(k) = {} & \beta V(k-d)+\Psi^n(k)^2+2\Psi^n(k)e^n(k) \\ & +\phi^n(k-d)^t F(k)\phi^n(k-d)e^n(k)^2 \end{aligned} \tag{5.99}$$

Multiplying (5.99) through by a^{2k} and making use of the exponentially weighted signals defined in (5.96) results in

$$\begin{aligned} & a^{2k}[V(k)-\beta V(k-d)] \\ & \quad =\bar{\Psi}^n(k)^2+2\bar{\Psi}^n(k)\bar{e}^n(k)+\phi^n(k-d)^t F(k)\phi^n(k-d)\bar{e}^n(k)^2 \end{aligned} \tag{5.100}$$

Define

$$V'(k) = \tilde{\theta}(k)^t F'(k)^{-1}\tilde{\theta}(k) \tag{5.101}$$

where $F'(k)^{-1}=a^{2k}\cdot F(k)^{-1}$. Therefore from (5.98) and (5.101)

$$V'(k) = a^{2k}\cdot V(k) \tag{5.102}$$

and

$$a^{2d}\cdot V'(k-d) = a^{2k}\cdot V(k-d) \tag{5.103}$$

Substituting (5.102) and (5.103) into (5.100) yields

$$V'(k)-\beta\alpha^{2d}V'(k-d)=\Psi^n(k)^2+2\Psi^n(k)\bar{e}^n(k)$$
$$+\phi^n(k-d)^tF(k)\phi^n(k-d)\bar{e}^n(k)^2 \qquad (5.104)$$
$$=V'(k)-V'(k-d)+(1-\beta\alpha^{2d})V'(k-d)$$

Since $(1-\beta\alpha^{2d})\geq 0$, then

$$\sum_{k=0}^{N}[\Psi^n(k)^2+2\bar{\Psi}^n(k)\bar{e}^n(k)+\phi^n(k-d)^tF(k)\phi^n(k-d)\bar{e}^n(k)^2]$$
$$\geq -\sum_{k=-1}^{-d} V'(k) \qquad (5.105)$$

From Lemma 5.1 and using the result that $|\phi^n(k-d)|^2<1$, it may be concluded that

$$\lambda_1/(\beta+\lambda_1) \geq \phi^n(k-d)^tF(k)\phi^n(k-d) \qquad (5.106)$$

which completes the proof of this lemma.

As remarked by Lim (1982), if β in (5.24) is chosen strictly less than unity then there exists an $\alpha>1$ such that $(1-\beta\alpha^{2d})\geq 0$. If $\alpha<1$ then $(1-\beta\alpha^{2d})\geq 0$ for $0<\beta\leq 1$. However, if $\beta<1$, then some difficulty arises with guaranteeing boundedness of λ_1, the maximum eigenvalue of F(k). From (5.24), it is clear that when the plant is at steady-state (i.e. $\phi^n(k-d)\approx 0$) the eigenvalues of F(k) will be unbounded (i.e. $\lambda_1\rightarrow\infty$). This will bring about the vanishing radius problem (i.e. $\sigma_{LS}\rightarrow 1$) as discussed in section 5.3.

One way to avoid the vanishing radius problem in this case is to use an algorithm which forces an upper bound on λ_1. For example the RLS scheme in (5.58) and (5.59) used by Ortega *et al* (1985) ensures that the eigenvalues of F(k) are contained in the user selected interval $[\lambda_0,\lambda_1]$. Ortega *et al* (1985) prove that the exponentially weighted counterpart $\hat{H}_1{}^n$ of $H_1{}^n:e^n(k)\rightarrow\Psi^n(k)$ defined by (5.58) and

(5.59) is outside the cone$(-1,(1-\sigma_{RLS})^{1/2})$ for α and σ_{RLS} satisfying certain inequalities. In the proof of this result, Ortega *et al* (1985) define the matrix $F_1(k)$, and the quadratic functions $V(k)$ and $V_1(k)$ as

$$F_1(k) = F(k-d)-\frac{F(k-d)\phi^n(k-d)\phi^n(k-d)^tF(k-d)}{\lambda+\phi^n(k-d)^tF(k-d)\phi^n(k-d)} \qquad (5.107)$$

$$V(k)=\tilde{\theta}(k)^tF(k)^{-1}\tilde{\theta}(k)/\lambda \qquad (5.108)$$

$$V_1(k)=\tilde{\theta}(k)^tF_1(k)^{-1}\tilde{\theta}(k)/\lambda \qquad (5.109)$$

Ortega *et al* (1985) claim that after some algebra it may be shown that

$$V_1(k)-V(k-d)=\Psi^n(k)^2+2\Psi^n(k)e^n(k)+[\sigma(k)/(\lambda+\sigma(k))]e^n(k)^2 \qquad (5.110)$$

where $\sigma(k)$ has been used to replace $\phi^n(k-d)^tF(k-d)\phi^n(k-d)$ for brevity. However direct substitution of (5.58) and (5.59) into (5.109) gives

$$\begin{aligned} V_1(k)-V(k-d)=&\Psi^n(k)^2/\lambda^2 + (2/\lambda)(1-\lambda_0/\lambda_1)\Psi^n(k)e^n(k) \\ &+ (1-\lambda_0/\lambda_1)^2[\sigma(k)/(\lambda+\sigma(k))]e^n(k)^2 \\ &+ (2/\lambda)\lambda_0\tilde{\theta}(k-d)^tF_1(k)^{-1}\phi^n(k-d)e^n(k) \\ &+ (2\lambda_0/\lambda\lambda_1)(\lambda_1-\lambda_0)\phi^n(k-d)^t\phi^n(k-d)e^n(k)^2 \\ &+ (\lambda_0{}^2/\lambda)\phi^n(k-d)^tF_1(k)^{-1}\phi^n(k-d)e^n(k)^2 \end{aligned} \qquad (5.111)$$

The first three terms on the righthand side of (5.111) are similar to the righthand side of (5.110) but it is clear that (5.110) and (5.111) are not equivalent. If λ_0 is set equal to zero then (5.111) becomes

$$\begin{aligned} V_1(k)-V(k-d) = &\Psi^n(k)^2/\lambda^2+(2/\lambda)\Psi^n(k)e^n(k) \\ &+[\sigma(k)/(\lambda+\sigma(k))]e^n(k)^2 \end{aligned} \qquad (5.112)$$

However Ortega *et al* (1985) state that λ_0 is a strictly positive scalar and therefore (5.110) is incorrect. This error was also verified by a simple numerical example. As a result of this error the correctness of the input-output properties for $\hat{H}_1{}^n$ derived by Ortega *et al* (1985) is in question. Setting $\lambda_0=0$ in (5.59) gives (5.60). The dissipative property of the relation $H_1{}^n$ based on the adaptation mechanism (5.58) and (5.60) was derived in Lemma 5.4. The following lemma summarizes the properties of its exponentially weighted counterpart.

Lemma 5.7: Consider the relation $H_1{}^n:e^n(k)\rightarrow\Psi^n(k)$ defined by (5.58) and (5.60). Define its exponentially weighted counterpart as $\hat{H}_1{}^n:\bar{e}^n(k)\rightarrow\bar{\Psi}^n(k)$. $\hat{H}_1{}^n$ is weakly $(1,\lambda,\lambda^2\cdot\sigma_{RLS})$ dissipative for σ_{RLS} satisfying (5.61) if $(1-\alpha^{2d})\geq 0$.

Proof: Multiplying (5.64) through by α^{2k} gives

$$\alpha^{2k}[V(k)-V(k-d)] = (1/\lambda)\bar{\Psi}^n(k)^2+2\bar{\Psi}^n(k)\bar{e}^n(k) +\phi^n(k-d)^tF(k)\phi^n(k-d)\bar{e}^n(k)^2 \qquad (5.113)$$

Define V'(k) as in (5.101). Then (5.113) may be written as

$$V'(k)-\alpha^{2d}V'(k-d) = (1/\lambda)\bar{\Psi}^n(k)^2+2\bar{\Psi}^n(k)\bar{e}^n(k) +\phi^n(k-d)^tF(k)\phi^n(k-d)\bar{e}^n(k)^2 \qquad (5.114)$$

As stated in Lemma 5.4

$$\lambda\cdot\lambda_1/(\lambda+\lambda_1) > \phi^n(k-d)^tF(k)\phi^n(k-d) \qquad (5.115)$$

Since $(1-\alpha^{2d})\geq 0$, then

$$\sum_{k=0}^{N}[\bar{\Psi}^n(k)^2+2\lambda\bar{\Psi}^n(k)\bar{e}^n(k)+(\lambda^2\cdot\lambda_1/(\lambda+\lambda_1))\bar{e}^n(k)^2] \geq -\sum_{k=-1}^{-d} V'(k) \qquad (5.116)$$

which completes the proof of this lemma.

If $a<1$ then $(1-a^{2d})\geq 0$. However if $a>1$ then $(1-a^{2d})$ cannot be greater than or equal to zero.

The L_∞ extension of the previous L_2 results as presented by Ortega *et al* (1985) is rederived below to emphasize the role of the exponentially weighted signals in obtaining a L_∞ result.

Lemma 5.8(see also Lemma 4.2 of Ortega *et al* (1985)): Consider the error feedback system defined in (5.54) with the least squares adaptation algorithm presented in (5.23) and (5.24). If

$\bar{H}_2{}^n = a^k \bar{H}_2 a^{-k}$ is strictly inside the

cone$(\sigma_{LS}{}^{-1}, \sigma_{LS}{}^{-1}(1-\sigma_{LS})^{1/2})$ with $a>1$ and $(1-\beta a^{2d})\geq 0$,

then $\Psi^n(k)\epsilon L_\infty$.

Proof: L_2 stability of map $\bar{e}^n(k)^* \rightarrow \Psi^n(k)$ is guaranteed from Lemma 5.6 and Theorem 5.1, i.e. there exists a $K<\infty$ such that

$$||\Psi^n(k)||_N{}^2 \leq K\cdot||\bar{e}^n(k)^*||_N{}^2 \quad \forall N \tag{5.117}$$

By definition

$$||\Psi^n(k)||_N{}^2 = \sum_{k=0}^{N}(\Psi^n(k))^2 \tag{5.118}$$

$$= \sum_{k=0}^{N}(a^k\psi^n(k))^2$$

$$\geq (a^N\psi^n(N))^2 \quad \text{(last term in series)} \tag{5.119}$$

and

$$||\bar{e}^n(k)^*||_N{}^2 = \sum_{k=0}^{N}(\bar{e}^n(k)^*)^2 \tag{5.120}$$

$$= \sum_{k=0}^{N}(a^k e^n(k)^*)^2$$

$$\leq ||e^n(k)^{*2}||_\infty \cdot \sum_{k=0}^{N} a^{2k} \tag{5.121}$$

Since $a>1$

$$\sum_{k=0}^{N} \alpha^{2k} = \frac{\alpha^{2N}-\alpha^{-2}}{1-\alpha^{-2}} \le \frac{\alpha^{2N}}{1-\alpha^{-2}} \qquad (5.122)$$

From combining (5.117), (5.119), (5.121) and (5.122) it may be concluded that

$$\Psi^{n}(N)^{2} \le \frac{K}{\bar{\rho}(1-\alpha^{-2})} \cdot ||e(k)^{*2}||_{\infty} \qquad (5.123)$$

where $\bar{\rho} \le \rho(k)\ \forall k$ from (5.19). This completes the proof of Lemma 5.8.

The importance of α being greater than unity in the proof of Lemma 5.8 is evident from (5.122). If $\alpha<1$, then the upper bound in (5.122) is no longer valid, i.e.

$$\frac{\alpha^{2N}-\alpha^{-2}}{1-\alpha^{-2}} \not\le \frac{\alpha^{2N}}{1-\alpha^{-2}} \qquad (5.124)$$

For this case, combining (5.117), (5.119), (5.121) and (5.124) gives

$$\alpha^{2N}\Psi^{n}(N)^{2} \le \frac{K}{\bar{\rho}} \cdot \left(\frac{1-\alpha^{2N+2}}{1-\alpha^{2}}\right) \cdot ||e(k)^{*2}||_{\infty} \qquad (5.125)$$

and multiplying (5.125) through by α^{-2N} results in

$$\Psi^{n}(N)^{2} \le \frac{K}{\bar{\rho}} \cdot \left(\frac{\alpha^{-2N}-\alpha^{2}}{1-\alpha^{2}}\right) \cdot ||e(k)^{*2}||_{\infty} \qquad (5.126)$$

In this case with $\alpha<1$

$$\alpha^{-2N} \to \infty \text{ as } N \to \infty \qquad (5.127)$$

Therefore the upper bound on $\Psi^{n}(k)$ in (5.126) is unbounded and hence it cannot be concluded that $\Psi^{n}(k) \in L_{\infty}$.

As may be seen from the analysis in Lemma 5.8, exponential weighting allows for the derivation of a L_{∞}

result from a L_2 result. However the value of α must be selected strictly greater than unity in order to achieve this result. This restriction on α in turn limits the allowable parameter adaptation algorithms which may be considered for the L_∞ case. For example, the RLS scheme examined in Lemma 5.7 does not qualify because the input-output properties cannot be derived with $\alpha>1$. On the other hand, the dissipative property of the least squares algorithm considered with the forgetting factor β in Lemma 5.6 may be derived with $\alpha>1$.

Ortega *et al* (1985) did not carry out their L_∞ analysis on the constant gain algorithm in (5.55). They did not make it clear in their paper if the analysis could not be performed on this particular algorithm or whether the L_∞ extension using the constant gain algorithm is straightforward but was omitted strictly for brevity purposes. In the following lemma, the input-output properties of the exponentially weighted operator $\bar{H}_1{}^n$ are derived for the constant gain algorithm.

Lemma 5.9: Consider the relation $H_1{}^n: e^n(k) \to \Psi^n(k)$ defined by (5.55). Define its exponentially weighted counterpart as $\bar{H}_1{}^n: \bar{e}^n(k) \to \bar{\Psi}^n(k)$. $(\bar{H}_1{}^n + \sigma_{CG}/2)$ is passive for all $\sigma_{CG} > f$ if $\alpha^{2d} < 1$.

Proof: Consider the quadratic function

$$V(k) = \tilde{\theta}(k)^t f^{-1} \tilde{\theta}(k)/2 \qquad (5.128)$$

Substituting (5.55) into (5.128) gives

$$V(k) = V(k-d)+\Psi^n(k)e^n(k)+\phi^n(k-d)^t f\phi^n(k-d)e^n(k)^2/2 \tag{5.129}$$

Multiplying (5.129) through by α^{2k} yields

$$\alpha^{2k}[V(k)-V(k-d)] = \Psi^n(k)\bar{e}^n(k) + \phi^n(k-d)^t f\phi^n(k-d)\bar{e}^n(k)^2/2 \tag{5.130}$$

Summing (5.130) from 0 to N leads to

$$\sum_{k=0}^{N} \alpha^{2k}[V(k)-V(k-d)] = (1-\alpha^{2d})\cdot\sum_{k=0}^{N-d} V(k)\alpha^{2k} + \sum_{k=N-d+1}^{N} V(k)\alpha^{2k} - \sum_{k=-1}^{-d} \alpha^{2(k+d)}V(k) \tag{5.131}$$

Since $\alpha^{2d}<1$, then

$$\sum_{k=0}^{N}[\Psi^n(k)\bar{e}^n(k)+\sigma_{CG}\bar{e}^n(k)^2/2] \geq -\sum_{k=-1}^{-d} \alpha^{2(k+d)}V(k) \tag{5.132}$$

where $\sigma_{CG}>f\geq\phi^n(k-d)^t f\phi^n(k-d)$. This completes the proof of this lemma.

If $\alpha^{2d}<1$ then $\alpha<1$. The next step would be to continue with the L_∞ extension as done in Lemma 5.8 for the least squares algorithm. However as demonstrated in the discussion immediately following Lemma 5.8, if $\alpha<1$ it cannot be shown that $\Psi^n(k)\epsilon L_\infty$ and the L_∞ stability result does not follow. Therefore the constant gain algorithm does not qualify for the L_∞ case which perhaps answers the question why Ortega *et al* (1985) did not extend their results in this direction.

Of the three algorithms considered in Lemmas 5.6, 5.7 and 5.9, only the least squares algorithm with the forgetting factor β in (5.23) and (5.24) is eligible for the L_∞ stability analysis. However as mentioned earlier this algorithm may result in a vanishing radius (i.e. $\sigma_{LS}\to 1$ in (5.97)) because the maximum eigenvalue of the gain matrix $F(k)$ may be unbounded (i.e. $\lambda_1\to\infty$). From inspection of (5.24) the difficulty arises when the normalized regressor becomes small which causes $F(k)^{-1}$ to decay to zero (or $F(k)$ to become large). In order to prevent a vanishing radius, the algorithm must be modified in such a way as to impose limits on the eigenvalues of $F(k)$. This may be accomplished via a covariance resetting feature based on a check of the eigenvalues of $F(k)$ at each sampling instant. For instance, if at some instant k, $\lambda_{max}F(k)$ exceeds some designer selected bound, λ_u, then $F(k)$ would be reset such that $\lambda_{max}F(k)<\lambda_u$. This feature remains to be formally included in the derivation of the input-output properties of $\bar{H}_1^n$ in Lemma 5.6.

The next step in the L_∞ analysis is to derive conditions on H_2 which ensure that the conditions on $\bar{H}_2^n$ in Lemma 5.8 are met. Ortega *et al* (1985) have stated in their paper that the conditions derived for the L_2 case in Lemma 5.5 (i.e. with respect to $H[(\mu^{1/2}q)^{-1}]$) carry over directly for the L_∞ case. In the following lemma it is demonstrated that their statement is strictly not correct and needs to be replaced by a condition on $H[(\mu q)^{-1}]$.

Lemma 5.10: Consider the relation $H:x(k)\to y(k)$ where H is a linear, time-invariant rational function in the backward shift operator q^{-1}. If $H[(\mu q)^{-1}]$ is inside the cone(C,R) and if $\alpha^2=\mu^{-1}$, then $\bar{H}^n:\bar{x}^n(k)\to\bar{y}^n(k)$ is inside the same cone(C,R).

Proof: Define

$$w_2(k) = (\bar{y}^n(k)-C\bar{x}^n(k))^2 - (R\bar{x}^n(k))^2 \tag{5.133}$$

$$= \alpha^{2k}\rho(k)^{-1}w(k)$$

where w(k) was defined in (5.75). Taking the sum from 0 to N of $w_2(k)$ gives

$$\sum_{k=0}^{N} w_2(k) = \sum_{k=0}^{N}\mu^{-k}\rho(k)^{-1}w(k) \tag{5.134}$$

$$= \sum_{k=0}^{N}\mu^{k}\rho(k)^{-1}\mu^{-2k}w(k)$$

$$= \mu^{N+1}\rho(N+1)^{-1}\cdot\sum_{k=0}^{N}\mu^{-2k}w(k) \tag{5.135}$$

$$- \sum_{j=0}^{N}\Big[\Big(\sum_{k=0}^{j}\mu^{-2k}w(k)\Big)\cdot\big(\mu^{j+1}\rho(j+1)^{-1}-\mu^{j}\rho(j)^{-1}\big)\Big]$$

(5.135) may be verified by straightforward expansion. It was shown in Lemma 5.5 that the normalization factor in (5.19) verifies

$$\mu^{j+1}\rho(j+1)^{-1} - \mu^{j}\rho(j)^{-1} < 0 \quad \forall j \tag{5.136}$$

Expressing $H(q^{-1})$ as in (5.81) and (5.82) and multiplying (5.82) through by μ^{-k} gives

$$b_0\mu^{-k}y(k)+\cdots+b_m\mu^{-k}y(k-m)$$

$$= a_0\mu^{-k}x(k)+\cdots+a_n\mu^{-k}x(k-n) \tag{5.137}$$

Define

$$x''(k) = \mu^{-k}x(k) \tag{5.138}$$

$$y''(k) = \mu^{-k}y(k) \tag{5.139}$$

Substituting (5.138) and (5.139) into (5.137) yields

$$b_0y''(k)+b_1\mu^{-1}y''(k-1)+\cdots+b_m\mu^{-m}y''(k-m)$$

$$= a_0x''(k)+\cdots+a_n\mu^{-n}x''(k-n) \tag{5.140}$$

which may be expressed as

$$y''(k) = H[(\mu q)^{-1}]\cdot x''(k) \tag{5.141}$$

Therefore if $H[(\mu q)^{-1}]$ is inside the cone(C,R), then

$$\begin{aligned}&\sum_{k=0}^{N}[(y''(k)-Cx''(k))^2-(Rx''(k))^2]\\ &\quad=\sum_{k=0}^{N}[(\mu^{-k}y(k)-C\mu^{-k}x(k))^2-(R\mu^{-k}x(k))^2]\\ &\quad\leq 0\end{aligned} \tag{5.142}$$

which implies that

$$\sum_{k=0}^{N}\mu^{-2k}w(k)\leq 0 \tag{5.143}$$

From combining (5.135), (5.136) and (5.143) it may be concluded that

$$\sum_{k=0}^{N}w_2(k)\leq 0 \tag{5.144}$$

Therefore $\tilde{H}^n$ is also inside the cone(C,R) which completes the proof of this lemma.

From Lemma 5.8 it is clear that α must be chosen strictly greater than unity. This condition is satisfied in Lemma 5.10 where $\alpha^2=\mu^{-1}$. From Lemma 5.6, the dissipative properties for the least squares algorithm in (5.23) and (5.24) may be derived if $(1-\beta\alpha^{2d})\geq 0$. Combining these restrictions on α leads to the condition

$$\beta\leq\mu^d \tag{5.145}$$

Therefore from (5.145) it may be concluded that β in (5.24) and μ in (5.19) cannot be selected independent of each other but must satisfy (5.145) for this analysis to follow.

5.6 Addition of P and Q Weighting to the Control Law

The majority of the discussion to this point in this chapter has dealt with aspects of the parameter adaptation algorithms. The other key element of an adaptive control system is the control law itself. One of the important features of Lim's (1982) work was that his robustness analysis incorporated weighting polynomials in the control law which are characteristic of the Clarke and Gawthrop (1975) self-tuning controller. The analysis performed by Ortega *et al* (1985) was based on a pole-placement, all-zero cancelling control law. Although Ortega *et al* (1985) removed the signal boundedness assumption discussed in section 5.3, their control law did not include the same weighting polynomials found in Lim's (1982) work. The objective of this section is to extend Ortega *et al*'s (1985) controller structure to include the same weighting polynomials considered by Lim (1982).

Assume that the plant may be described by (5.47) (set $\xi(k)=0$ for clarity). Define an augmented plant with output z, and input u, where

$$z(k) = P(q^{-1})y(k) + q^{-d}Q(q^{-1})u(k) \tag{5.146}$$

and P and Q are weighting polynomials in the delay operator q^{-1}. Substituting (5.47) into (5.146) gives

$$Az(k) = q^{-d}(PB + QA)u(k) \tag{5.147}$$

and by defining $B'=PB+QA$, (5.147) becomes

$$Az(k) = q^{-d}B'u(k) \tag{5.148}$$

The tracking error is defined as before in (5.48) with the filtered output $C_Ry(k)$ replaced by z(k), i.e.

$$e(k) = z(k)-r(k) \tag{5.149}$$

The regulator structure for the augmented system is based on a predictive control law as in (5.49) and is given by

$$\begin{aligned} r(k+d) &= \hat{S}(k)\,u(k) + \hat{R}(k)\,z(k) \qquad (5.150)\\ &= \hat{\theta}(k)^t\phi(k) \end{aligned}$$

where

$$\phi(k)^t = [u(k)\cdots u(k-n_s), z(k)\cdots z(k-n_r)] \tag{5.151}$$

The stabilizability assumption is introduced as before and the polynomial C is defined in terms of tuned parameters $\theta^{*t}=[S^*\ R^*]$ where

$$C = AS^*+q^{-d}R^*B' \tag{5.152}$$

Adding $q^{-d}R^*B'z(k)$ to both sides of (5.148) multiplied by S^* gives

$$AS^*z(k) + q^{-d}R^*B'z(k) = B'S^*u(k-d) + B'R^*z(k-d) \tag{5.153}$$

Substituting (5.152) into (5.153) and rearranging

$$z(k) = \frac{B'}{C}\cdot[S^*u(k-d)+R^*z(k-d)] \tag{5.154}$$

$$= H_2\cdot\theta^{*t}\phi(k-d) \tag{5.155}$$

where $H_2=C^{-1}B'$. From (5.155), an error feedback system based on this augmented system may be derived, i.e.

$$\begin{aligned} e(k) &= z(k)-r(k)\\ &= -H_2(r(k)-\theta^{*t}\phi(k-d)) + H_2r(k) - r(k) \qquad (5.156)\\ &= -H_2(\hat{\theta}(k-d)^t\phi(k-d)-\theta^{*t}\phi(k-d))+(H_2-1)r(k) \qquad (5.157)\\ e(k) &= -H_2\Psi(k) + e(k)^* \qquad (5.158) \end{aligned}$$

and

$$\Psi(k) = H_1e(k) \tag{5.159}$$

where $e(k)^*=(H_2-1)r(k)$ and $\Psi(k)$ is defined in (5.1) with $\theta^*=\theta$. (5.158) and (5.159) now define an error system which extends the system in (5.54) to include P and Q weighting polynomials in the predictive control law. The stability analysis based on input-output theory may now proceed from this point where the adaptation algorithms discussed in the earlier sections would now be used to estimate the controller parameters in (5.150).

5.7 Conclusions

1. The least squares type of parameter estimation algorithms used by both Lim (1982) and Ortega *et al* (1985) do not satisfy a conic sector property but instead satisfy a weakly dissipative property. In the case of the sector condition, an incorrect *a priori* assumption of zero initial parameter error was made. A modified form of the conic sector theorem is developed which accomodates the nonzero initial parameter errors.

2. An *a priori* signal boundedness assumption was used throughout much of the early literature dealing with the input-output stability approach to the analysis of adaptive systems. Removal of this assumption prevents proof of stability of adaptive control systems based on least squares parameter estimators even when no unmodeled dynamics are present due to the vanishing radius problem. One of the key contributions of Ortega *et al* (1985) was the inclusion of normalized signals in the parameter estimation scheme which prevents the vanishing radius and allows the stability analysis to proceed.

3. It is demonstrated that the Key Technical Lemma (KTL) of Goodwin *et al* (1980) is one means of proving L_2 stability of unnormalized signals from the L_2 stability of the corresponding normalized signals. Also an alternate proof of stability for projection algorithm I of Goodwin *et al* (1980) is presented based on the modified sector theorem to illustrate this role of the KTL. Some explanation for the specific choice of normalization factor made by Ortega *et al* (1985) is given. It turns out that their normalization factor allows for a straightforward derivation of the sector conditions on the linear, time-invariant operator which contains information about the unmodeled dynamics.

4. It is demonstrated how L_2 stability results may be extended to the L_∞ case. The input-output properties of three different parameter estimation schemes are derived. The need for exponentially weighted signals to obtain a L_∞ result is illustrated. It turns out that of the three algorithms examined, only the least squares algorithm with a forgetting factor satisfies the condition required to extend the L_2 results.

5. The predictive control law used by Ortega *et al* (1985) has been augmented to include weighting polynomials on the plant input and output. These weighting polynomials, which were originally considered by Lim (1982) in his analysis, may now be included in the conic sector formulation.

5.8 References

Clarke, D.W. and P.J. Gawthrop (1975). Self-tuning controller. *Proc. IEE*, 122, 929-934.

Egardt, B. (1979). *Stability of Adaptive Controllers*. Springer-Verlag, Berlin.

Gawthrop, P.J. (1980). On the stability and convergence of a self-tuning controller. *Int. J. Contr.*, 31, 973-998.

Gawthrop, P.J. and K.W. Lim (1982). Robustness of self-tuning controllers. *IEE Proc.*, 129, 21-29.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Hill, D.J. and P.J. Moylan (1983). General instability results for interconnected systems. *SIAM J. Contr. Optimiz.*, 21, 256-278.

Lim, K.W. (1982). Robustness of self-tuning controllers. Dept. of Eng. Sci., Oxford University, OUEL Report No. 1422/82.

Ortega, R. and I.D. Landau (1983a). On the model-process mismatch tolerance of various parameter adaptation algorithms in direct control schemes: a sectoricity approach. *IFAC Workshop Adapt. Syst.*, San Francisco, CA.

Ortega, R. and I.D. Landau (1983b). On the design of robustly performing adaptive controllers for partially modelled systems. *Proc. 22nd IEEE CDC*, 967-971.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Safanov, M.G. (1980). *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, MA.

6. A New Global Stability Result for the Model-Plant Mismatch Problem

6.1 Introduction

Most stability results for adaptive control systems are based on the assumption that the order of the model used in the control structure is greater than or equal to that of the process. However most real processes are high order and hence, in general, a lower order control structure is used in practice. Rohrs (1982) demonstrated that straightforward application of the 'stable' algorithms found in the literature may lead to stability problems when unmodeled dynamics, due to this model-process order mismatch, are present. Results such as those of Rohrs (1982) have sparked a new wave of research into the performance and behaviour of adaptive control systems when the model order assumption is violated.

Kosut and Johnson (1984) suggested that the main reason for the lack of robust adaptive control theory is that the emphasis in the early stability studies was on *global* results. What is meant by 'global' in this context is that the objective is to require as little *a priori* information about process parameters, nature of external signals, initial parameter error, etc. as possible to prove

---

A version of this chapter has been submitted for publication. Cluett, W.R., J.M. Martin-Sanchez, S.L. Shah and D.G. Fisher (1986). IEEE Transactions on Automatic Control. Another version of this chapter has been accepted for publication. Cluett, W.R., S.L. Shah and D.G. Fisher (1986). International Journal of Control.

stability. Kosut and Johnson (1984), in their stability analysis of adaptive systems in the presence of unmodeled dynamics, assume that there exist sets of *tuned* controller parameters associated with particular sets of external signals. Each tuned set is defined such that the process would be stabilized by a fixed parameter controller containing the tuned parameters. The union of these tuned sets is then used to develop a generic adaptive error system in which a tuned interconnection operator must satisfy a strictly positive real (SPR) condition in order to obtain a global stability result. These results are limited by this last SPR condition which is almost surely violated due to typically unmodeled, high frequency dynamics. This conclusion is the justification given by Kosut and Johnson (1984) to consider conditions for *local* rather than global stability. In this context 'local' means that restrictions are placed on the nature of the external signals and on the initial adaptive parameter error in order to guarantee stability.

Analogous results in the discrete-time case were presented by Ortega *et al* (1985). Conic sector stability theory (Safanov, 1980) was used to analyze the corresponding generic error system. The discrete operator in the error system which is a function of the unmodeled dynamics, and is also obtained by assuming the existence of a tuned set of parameters, is restricted by a conicity condition which corresponds to the SPR condition imposed on its continuous-time counterpart. Therefore this result would appear to be as restrictive a condition in the discrete case as it is for the continuous case and hence would also suggest a redirection away from the global stability approach.

The results obtained by Cluett *et al* (1986) indicate that the destablilizing effect of high frequency unmodeled dynamics present in a continuous system are lessened by sampling to the extent that global stability results need not be abandoned. It is also demonstrated by Cluett *et al* (1986) that the sampling rates which bring about the necessary reduction are reasonable and within the recommended guidelines for the choice of sampling times in discrete control.

Recent results of Praly (1984) present an extension of those of Ortega *et al* (1985) in the sense that the analysis now allows for the consideration of L_∞ disturbances, time variations and nonlinearities in addition to unmodeled dynamics. However the basic assumptions, including that of the existence of a tuned set of parameters, and the conicity analysis remain basically the same as in Ortega *et al* (1985). A parameter projection strategy, used in Egardt (1979) to handle the problem of bounded disturbances, has been introduced and combined in Praly (1984) with a normalization procedure which causes the adaptation law to see the unmodeled effects as a bounded disturbance. This normalization procedure is assumed to provide boundedness of these unmodeled residuals, but this critical property is not proven.

Adaptive Predictive Control Systems (APCS) were designed to provide a basis for the formal proof of global stability for discrete-time systems (Martin-Sanchez, 1974, 1976a,b). Results obtained by Martin-Sanchez (1982), and Martin-Sanchez *et al* (1984) prove global stability of an APCS algorithm for a general class of time-invariant, minimum phase processes in the presence of bounded, unmeasured noise and disturbances. More recently in Martin-Sanchez (1985), conditions have been established for global stability of an APCS when applied to a general class of stochastic, time-varying, minimum phase processes.

The analysis presented by Martin-Sanchez (1982) and Martin-Sanchez *et al* (1984) defines a perturbation signal which includes the effect of noise and disturbances on the process output. This perturbation signal is a key factor in the introduction of an on/off criterion for parameter adaptation which is used in the proof of global stability. Martin-Sanchez *et al* (1984) also suggested that unmodeled dynamics may be included in the perturbation signal as long as they are bounded. However a general solution to the unmodeled dynamics problem cannot be expected to satisfy this restrictive assumption.

This chapter presents a general approach for the stability analysis of adaptive predictive control systems based on a main theorem which states a *single* condition for stability in terms of convergence properties of the *a posteriori* estimation error and the estimated parameters. This general approach has been used to prove global stability for a class of stable-inverse processes in the presence of L_∞ disturbances and unmodeled dynamics due to model-process order mismatch. In this result, a normalized estimation system is used which permits a formal proof that the modeling errors can be treated as a bounded disturbance. The on/off mechanism considered by Martin-Sanchez (1982) and Martin-Sanchez *et al* (1984) is therefore valid. This normalized estimation system verifies the convergence properties that guarantee global stability. In comparison with the work of Praly (1984), the type of normalization used here has the advantage of formally guaranteeing the boundedness of the unmodeled dynamics term and, in general, of the normalized perturbation signal. Furthermore, this solution is not restricted by any SPR type of condition, nor does it assume the existence of a tuned set of parameters. In fact, it is proven that the adaptive control system presented here ensures parameter convergence to a tuned set. The global stability result appears as a natural conclusion rather than a difficult goal to attain.

In this chapter it is also demonstrated how different stability problems (deterministic, stochastic, time-varying and unmodeled dynamics) can be cast under the general formulation presented here. Simulation experiments, based on a benchmark example (Rohrs, 1982), illustrate the theoretical results.

Other results related to the work presented in this chapter have recently appeared. Kreisselmeier and Anderson (1986) have used a relative error signal with a dead zone and a projection in the adaptive law. Ioannou and Tsakalis (1985) have included a retardation term and a normalizing signal in the adaptive law.

6.2 General Conditions for Stability

6.2.1 Process Description

Let the process be described by the following discrete, SISO, time-variant ARMA model

$$y_a(k) = \theta(k)^t \phi_a(k-d) + \xi(k) \tag{6.1}$$

where

$$\phi_a(k-d)^t = [y_a(k-d)\ y_a(k-d-1) \cdots u_a(k-d)\ u_a(k-d-1) \cdots]$$

is a vector of past true values of the process output, y_a, and input, u_a and d denotes the time delay of the process. (In previous APCS literature d has been represented by r+1, where r denotes the pure time delay of the process.) $\theta(k)$ is a time varying vector of the unknown process parameters. The dimension of ϕ_a and θ is determined by the process order. $\xi(k)$ represents the effect of unmeasured disturbances on the process output at time k.

The available measured process variables differ from the actual values due to measurement error, noise, etc., i.e.

$$y(k) = y_a(k) + n_y(k)$$

$$u(k) = u_a(k) + n_u(k) \tag{6.2}$$

and the corresponding measured ϕ vector becomes

$$\phi(k) = \phi_a(k) + n_\phi(k) \tag{6.3}$$

Substitution of (6.2) and (6.3) into the process equation (6.1) yields

$$y(k) = \theta(k)^t \phi(k-d) + \Delta(k) \tag{6.4}$$

where the perturbation signal, $\Delta(k)$, is given by

$$\Delta(k) = n_y(k) - \theta(k)^t n_\phi(k-d) + \xi(k) \tag{6.5}$$

6.2.2 APCS Description

The adaptive predictive (AP) model estimation, $\hat{y}(k|k)$, of the process output at time k is based on a vector of estimated parameters, $\hat{\theta}_r(k)$, and may be computed by

$$\hat{y}(k|k) = \hat{\theta}_r(k)^t \phi_r(k-d) \tag{6.6}$$

where

$$\phi_r(k-d)^t = [y(k-d)\ y(k-d-1) \cdots u(k-d)\ u(k-d-1) \cdots]$$

The dimension of ϕ_r and $\hat{\theta}_r$ is equal to or less than the dimension of ϕ and θ in (6.4). (The dimension of ϕ_r and $\hat{\theta}_r$ may also be chosen greater than that of ϕ and θ. In fact the dimension of ϕ and θ may be considered as large as desired by adding zero parameters to the θ vector.) $\phi_r(k-d)$ contains a subset of the more recent process inputs and outputs in $\phi(k-d)$. The corresponding *a posteriori* error is

$$e(k|k) = y(k) - \hat{y}(k|k) = y(k) - \hat{\theta}_r(k)^t \phi_r(k-d) \tag{6.7}$$

where the estimated parameter vector $\hat{\theta}_r(k)$ is generated by an adaptive law from the information available from the process at sampling time k.

Using APCS predictive control, the control input u(k) is computed to make the predicted output at time k+d equal to the desired output at the same instant, i.e.

$$y_d(k+d) = \hat{\theta}_r(k)^t \phi_r(k) \tag{6.8}$$

In order to solve for u(k) from (6.8) it must be shown that the adaptive law of the AP model parameters will always produce a nonzero value for the leading coefficient associated with u(k). The tracking error is defined by

$$\epsilon(k) = y(k) - y_d(k) \tag{6.9}$$

6.2.3 Basic Assumptions

The assumptions used to establish the stability conditions of APCS are presented here.

Assumption 1

The delay d is known.

Assumption 2

The desired process output y_d at time k+d is known at time k and is bounded, i.e. $|y_d(k+d)| \leq \lambda^2 < \infty \ \forall k$.

Assumption 3

The sequence $\{||\Phi(k)||\}$ is unbounded only if there is a subsequence $\{k_s\}$ such that

(a) $\lim_{k_s \to \infty} ||\Phi(k_s-d)|| = \infty$, and

(b) $|y(k_s)| > a_1 ||\Phi(k_s-d)|| - a_2, \ \forall k_s$

where Φ is an I/O regressor vector of dimension m, greater than or equal to the dimension n_p of ϕ and contains all of the inputs and outputs included in ϕ, and

$$0 < a_1 < \infty \text{ and } 0 \leq a_2 < \infty.$$

The conditions in assumption 3 were derived by Martin-Sanchez (1982) using a standard result stated by Goodwin *et al* (1980) and follow from the fact that the process in (6.4) has a stable-inverse.

6.2.4 Global Stability Condition

Stability results from previous papers (Martin-Sanchez, 1982; Martin-Sanchez *et al*, 1984; Martin-Sanchez, 1985) have been refined down to the one concise condition in the following theorem. This condition is not problem specific and depends only on the convergence properties of the *a posteriori* estimation error and the estimated parameters.

Theorem 6.1: Under assumptions 1-3, adaptive predictive control (6.8) of the process described by (6.4) guarantees a bounded sequence $\{||\Phi(k)||\}$ if the adaptive law of the AP model parameters satisfies the following condition

$$a_1 > ||\hat{\theta}_r(k)-\hat{\theta}_r(k-d)|| + \frac{|e(k|k)|}{\max(||\Phi(k-d)||,c)}$$

$$\forall k > k_1 > 0$$

where $k_1<\infty$ and c is any finite positive constant.

Proof: By combining (6.8) and (6.9), the tracking error may be rewritten as

$$\epsilon(k) = y(k) - \hat{\theta}_r(k-d)^t\phi_r(k-d) \qquad (6.10)$$

Subtracting (6.7) from (6.10)

$$\epsilon(k) = [\hat{\theta}_r(k)-\hat{\theta}_r(k-d)]^t\phi_r(k-d) + e(k|k) \tag{6.11}$$

Using the triangle and Cauchy-Schwarz inequalities

$$|\epsilon(k)| \leq ||\hat{\theta}_r(k)-\hat{\theta}_r(k-d)||\cdot||\phi_r(k-d)|| + |e(k|k)| \tag{6.12}$$

and using the fact that, by definition,

$||\Phi(k-d)||\geq||\phi_r(k-d)||$,

$$|\epsilon(k)| \leq ||\hat{\theta}_r(k)-\hat{\theta}_r(k-d)||\cdot||\Phi(k-d)|| + |e(k|k)| \tag{6.13}$$

From the definition of the tracking error and assumption 2

$$|\epsilon(k)| = |y(k)-y_d(k)| \geq |y(k)|-\lambda^2 \tag{6.14}$$

Combining (6.13) and (6.14)

$$\begin{aligned}|y(k)| \leq{} & ||\hat{\theta}_r(k)-\hat{\theta}_r(k-d)||\cdot||\Phi(k-d)|| \\ & + |e(k|k)| + \lambda^2\end{aligned} \tag{6.15}$$

Assume that the sequence $\{||\Phi(k)||\}$ is unbounded. Then using property (b) of assumption 3 and (6.15) gives

$$\begin{aligned}\alpha_1||\Phi(k_s-d)|| < {} & ||\hat{\theta}_r(k_s)-\hat{\theta}_r(k_s-d)||\cdot||\Phi(k_s-d)|| \\ & + |e(k_s|k_s)| + \lambda_1^2 \quad \forall k_s\end{aligned} \tag{6.16}$$

where $\lambda_1^2=\lambda^2+\alpha_2<\infty$.

Dividing both sides by $||\Phi(k_s-d)||$ we obtain

$$\alpha_1 < ||\hat{\theta}_r(k_s)-\hat{\theta}_r(k_s-d)||+\frac{|e(k_s|k_s)|}{||\Phi(k_s-d)||}+\frac{\lambda_1^2}{||\Phi(k_s-d)||} \tag{6.17}$$

Since $\lim\limits_{k_s\to\infty}||\Phi(k_s-d)||=\infty$, the last term on the right hand side of (6.17) tends to zero, and therefore (6.17) will violate the condition in Theorem 6.1. Therefore an unbounded sequence $\{||\Phi(k)||\}$ cannot exist and the theorem is proven.

6.2.5 A General Stability Result

In this section it will be demonstrated that the single condition for stability developed in Theorem 6.1 encompasses all of the previous APCS stability results and therefore shows that the approach used in this chapter is applicable to a broad class of adaptive systems. The types of systems to be analyzed are 1)deterministic, 2)stochastic, and 3)time-varying. The extension of this result to systems with unmodeled dynamics is the subject of section 6.3.

Deterministic Processes

A linear, time-invariant, SISO process for the deterministic or noise-free case may be obtained from the general process equation (6.4) by setting $\Delta(k)=0$ and $\theta(k)=\theta$,i.e.

$$y(k) = \theta^t\phi(k-d)$$

The APCS description follows as in subsection 6.2.2 except that the estimated parameter vector $\hat{\theta}_r(k)$ has dimension equal to the actual process parameter vector θ. It has been demonstrated by Martin-Sanchez (1982) and Martin-Sanchez *et al* (1984) that for a particular adaptive law the *a posteriori* error and the estimated parameters have the following properties:

(i) $\lim_{k\to\infty} e(k|k) = 0$

(ii) $\lim_{k\to\infty} [\hat{\theta}_r(k)-\hat{\theta}_r(k-d)] = 0$

It is clear that under properties (i) and (ii), the condition of Theorem 6.1 is satisfied for some finite k_1 and hence stability is guaranteed.

Stochastic Processes

A linear, time-invariant, SISO process for the stochastic case may be obtained from (6.4) by setting $\theta(k)=\theta$, i.e.

$$y(k) = \theta^t\phi(k-d) + \Delta(k)$$

As in the deterministic case it is assumed that the dimension of $\hat{\theta}_r$ is equal to that of θ. Martin-Sanchez (1982) and Martin-Sanchez *et al* (1984) show that for a particular adaptive law, the convergence properties of the estimation error and the estimated parameters are as follows:

(i) there exists k_2 such that $|e(k|k)|<2\Delta_b \quad \forall k>k_2$ where $|\Delta(k)|\leq\Delta_b<\infty \quad \forall k$

(ii) $\lim\limits_{k\to\infty} [\hat{\theta}_r(k)-\hat{\theta}_r(k-d)]=0$

Since c may be any finite positive constant, there exists a value for c such that

$$c\cdot\alpha_1 > 2\Delta_b$$

Therefore from properties (i) and (ii) a finite k_1 exists such that the condition of Theorem 6.1 is satisfied.

Time-Varying Processes

A linear, time-varying, SISO process may be described by (6.4). Some additional assumptions are imposed by Martin-Sanchez (1985) on the process in order to derive a set of convergence properties. These assumptions are as follows:

a) $[\sum_{k=1}^{\infty} \Delta(k)^2]^{1/2} < \infty$

b) The process parameters are bounded.

c) The number of changes in the process parameter values is finite.

Under this extended set of assumptions, the convergence properties of the *a posteriori* error and the estimated parameters for a particular adaptive law are:

(i) $\lim_{k\to\infty} e(k|k) = 0$

(ii) $\lim_{k\to\infty} [\hat{\theta}_r(k)-\hat{\theta}_r(k-d)] = 0$

As for the deterministic case, the above properties satisfy the condition of Theorem 6.1.

6.3 Verification of the Stability Condition in the Presence of Unmodeled Dynamics

In this section, a normalized parameter estimation system will first be defined. Then the convergence properties of the normalized *a posteriori* estimation error and the estimated parameters of this system will be proven. Finally, the properties of the *a posteriori* estimation error, considered in the previous section, necessary to guarantee global stability of the adaptive control scheme will be proven.

A SISO, time-invariant process will be used for further analysis and may be represented as

$$y(k) = \theta^t \phi(k-d) + \Delta(k) \tag{6.18}$$

where (6.18) is equivalent to (6.4) with $\theta(k)$ replaced by the time-invariant vector θ.

6.3.1 A Normalized Estimation System

The I/O variables of the process given by (6.18) in the normalized system will be defined in the following manner:

$$y^n(k) = y(k)/n(k)$$
$$x(k-d) = \phi(k-d)/n(k) \tag{6.19}$$
$$n(k) = \max(\max_{1\le i\le m} |\Phi_i(k-d)|, c)$$

where Φ_i is the ith element of Φ. Note that a suitable choice of the dimension m of Φ and the computation of a normalized factor such as n(k) are simple. Also note that the elements of the normalized regressor vector x are all contained in the interval [-1,1]. Using these normalized variables the process equation (6.18) may be rewritten as

$$y^n(k) = \theta^t x(k-d) + \Delta(k)/n(k) \tag{6.20}$$

In most cases the model order for adaptive control purposes is selected such that it is less than the actual process order. Again using the normalized variables defined in (6.19), the process representation (6.18) may be reexpressed in the form

$$y^n(k) = \theta_r{}^t x_r(k-d) + \theta_u{}^t x_u(k-d) + \Delta(k)/n(k) \tag{6.21}$$

where

$$\theta^t = [\theta_r{}^t \ \theta_u{}^t]$$
$$x(k-d)^t = [x_r(k-d)^t \ x_u(k-d)^t]$$

The dimension of θ_r and x_r is determined by the selected model order and is assumed to be equal to n_r. $x_u(k-d)$ is made up of the remaining terms in x(k-d) not included in $x_r(k-d)$ and is of dimension $n_p - n_r$.

If the last two terms on the right hand side of (6.21) are interpreted as a perturbation to the normalized process output $y^n(k)$, then (6.21) may be expressed in a form structurally similar to the original process representation in (6.18), i.e.

$$y^n(k) = \theta_r{}^t x_r(k-d) + \Delta^n(k) \tag{6.22}$$

where

$$\Delta^n(k) = \theta_u{}^t x_u(k-d) + \Delta(k)/n(k)$$

It is shown in the following lemma that the sequence $\{\Delta^n(k)\}$ defined in (6.22) is bounded.

Lemma 6.1: The normalized sequence of perturbation signals $\{\Delta^n(k)\}$ given by (6.22) is bounded if the sequence of perturbation signals $\{\Delta(k)\}$ given by (6.5) is bounded.

Proof: From the definition of $x_u(k-d)$ in (6.21) and of $n(k)$ in (6.19), it is clear that

$$||x_u(k-d)|| \le (n_p-n_r)^{1/2} \tag{6.23}$$

Using the Cauchy-Schwarz inequality gives

$$|\theta_u{}^t x_u(k-d)| \le (n_p-n_r)^{1/2} \cdot ||\theta_u|| \tag{6.24}$$

Since $||\theta_u||$ and $\{\Delta(k)\}$ are bounded, and $n(k) \ge c > 0$, then $\{\Delta^n(k)\}$ in (6.22) is bounded which completes the proof of this lemma.

Boundedness of $\{\Delta(k)\}$ requires only that the unmeasured noise and disturbances included in $\Delta(k)$ be bounded. This assumption is not restrictive and makes the result very practical and general. The type of normalization used here has the advantage of formally guaranteeing the boundedness of the unmodeled term in (6.22). The use of normalization to interpret the unmodeled effects as a disturbance was first proposed by Praly (1983). The normalization factor used by Praly (1983, 1984) was assumed to provide boundedness of the residuals but this critical property was not proven.

Lemma 6.1 proves boundedness of $\{\Delta^n(k)\}$ for mismatch between a linear model and a linear process. However the results of Praly (1984), and Kreisselmeier and Anderson (1986) deal with a wider class of unmodeled dynamics. In the result presented here, any *bounded* unmodeled process feature, e.g. some nonlinearities, may also be included in the normalized perturbation signal without further change.

The result of Lemma 6.1 is now used in the definition of the estimation system. The *a priori* estimation error for the normalized system in (6.22) is defined as

$$e^n(k|k-1) = y^n(k)-\hat{\theta}_r(k-1)^t x_r(k-d) \tag{6.25}$$

The estimated parameter vector is updated according to

$$\hat{\theta}_r(k) = \hat{\theta}_r(k-1)+\frac{\psi(k)^2 e^n(k|k-1)x_r(k-d)}{1+\psi(k)^2 x_r(k-d)^t x_r(k-d)} \tag{6.26}$$

The scalar $\psi(k)$ is determined by a criterion for stopping or continuing parameter adaptation.

(a) $\psi(k)^2=0$ if and only if

$$|e^n(k|k-1)|\le\Delta_b'(\psi_\ell,\Delta_b,k)<2\Delta_b<\infty \tag{6.27}$$

where

$$\Delta_b'(\psi_\ell,\Delta_b,k) = \frac{2+2\psi_\ell^2 x_r(k-d)^t x_r(k-d)}{2+\psi_\ell^2 x_r(k-d)^t x_r(k-d)}\cdot\Delta_b \tag{6.28}$$

with

$$0<\psi_\ell^2<\infty \text{ and } \Delta_b>\Delta_m=\max_{0<k\leq\infty} |\Delta^n(k)| \tag{6.29}$$

Δ_b is an estimate of an upper bound on the absolute value of the perturbation signal $\Delta^n(k)$ defined in (6.22), and Δ_m is the least upper bound.

(b) $\psi_\ell^2<\psi(k)^2\leq\psi_b(k)^2\leq\psi_u^2<\infty$ if and only if

$$|e^n(k|k-1)|>\Delta_b'(\psi_\ell,\Delta_b,k)\geq\Delta_b \tag{6.30}$$

where $\psi_b(k)^2$ is defined as follows:

$$\text{(i) } \psi_b(k)^2=\psi_u^2 \text{ if } |e^n(k|k-1)|>\Delta_b'(\psi_u,\Delta_b,k) \tag{6.31}$$

where according to (6.28):

$$\Delta_b'(\psi_u,\Delta_b,k) = \frac{2+2\psi_u^2 x_r(k-d)^t x_r(k-d)}{2+\psi_u^2 x_r(k-d)^t x_r(k-d)}\cdot\Delta_b \tag{6.32}$$

$$\text{(ii) } \psi_b(k)^2=\frac{2(|e^n(k|k-1)|-\Delta_b)}{(2\Delta_b-|e^n(k|k-1)|)x_r(k-d)^t x_r(k-d)} \tag{6.33}$$

if $\Delta_b'(\psi_\ell,\Delta_b,k)<|e^n(k|k-1)|\leq\Delta_b'(\psi_u,\Delta_b,k)$

Consequently, along the solution of the adaptive algorithm defined by (6.25)-(6.33), the adaptation of $\hat{\theta}_r(k)$ will be stopped at time k, i.e. $\hat{\theta}_r(k)=\hat{\theta}_r(k-1)$, if the absolute value of the normalized *a priori* estimation error, $|e^n(k|k-1)|$, is less than or equal to $\Delta_b'(\psi_\ell,\Delta_b,k)$. If the adaptation is not stopped, then the value of $\psi(k)^2$ is chosen in an interval greater than the selected value of ψ_ℓ^2 and less than or equal to a value $\psi_b(k)^2$. Appropriate use of this estimation system guarantees the nonsingularity of the leading coefficient associated with u(k) in the predictive control law (6.8). $\psi(k)^2$ is chosen from the allowed interval to ensure this property. This was proven by Martin-Sanchez *et al* (1984) where the unnormalized version of this algorithm was first presented.

6.3.2 Convergence Properties of the Normalized System

In this subsection under Lemmas 6.2, 6.3 and 6.4 we will prove certain convergence properties of the *a posteriori* estimation error of the normalized system, and of the estimated parameters. First, a relation between the *a priori* estimation error and the *a posteriori* error is presented.

The *a posteriori* estimation error for the normalized system in (6.22) is defined as

$$e^n(k|k) = y^n(k)-\hat{\theta}_r(k)^t x_r(k-d) \qquad (6.34)$$

Subtracting (6.34) from (6.25) gives

$$e^n(k|k-1) = [\hat{\theta}_r(k)-\hat{\theta}_r(k-1)]^t x_r(k-d)+e^n(k|k) \qquad (6.35)$$

Substituting (6.26) into (6.35) yields

$$e^n(k|k) = e^n(k|k-1)/(1+\psi(k)^2 x_r(k-d)^t x_r(k-d)) \qquad (6.36)$$

and the parameter update equation (6.26) may then be written as

$$\hat{\theta}_r(k) = \hat{\theta}_r(k-1)+\psi(k)^2 e^n(k|k) x_r(k-d) \qquad (6.37)$$

Lemma 6.2: Along the solution of the adaptive algorithm (6.25)-(6.33):

$$||\tilde{\theta}_r(k)||^2 - ||\tilde{\theta}_r(k-1)||^2 \le 0 \qquad \forall k$$

where $\tilde{\theta}_r(k)=\theta_r-\hat{\theta}_r(k)$.

Proof: Subtracting θ_r from both sides of (6.37) gives

$$\tilde{\theta}_r(k)+\psi(k)^2e^n(k|k)x_r(k-d) = \tilde{\theta}_r(k-1) \tag{6.38}$$

Hence

$$\begin{aligned}&||\tilde{\theta}_r(k)||^2+2\psi(k)^2e^n(k|k)\tilde{\theta}_r(k)^t x_r(k-d)\\&\quad+\psi(k)^4e^n(k|k)^2x_r(k-d)^t x_r(k-d) = ||\tilde{\theta}_r(k-1)||^2\end{aligned} \tag{6.39}$$

using (6.22) and (6.34), $e^n(k|k)$ may be expressed as

$$e^n(k|k) = \tilde{\theta}_r(k)^t x_r(k-d) + \Delta^n(k) \tag{6.40}$$

Combining (6.39) and (6.40) gives

$$\begin{aligned}||\tilde{\theta}_r(k)||^2-||\tilde{\theta}_r(k-1)||^2&=-2\psi(k)^2e^n(k|k)(e^n(k|k)-\Delta^n(k))\\&\quad-\psi(k)^4e^n(k|k)^2x_r(k-d)^t x_r(k-d)\end{aligned} \tag{6.41}$$

which may be rearranged as

$$\begin{aligned}&||\tilde{\theta}_r(k)||^2-||\tilde{\theta}_r(k-1)||^2\\&\quad=-\psi(k)^2e^n(k|k)^2(2+\psi(k)^2x_r(k-d)^t x_r(k-d))\\&\qquad+2\psi(k)^2e^n(k|k)\Delta^n(k)\end{aligned} \tag{6.42}$$

The left hand side of (6.42) will be equal to zero if $\psi(k)=0$ and less than or equal to zerc if $\psi(k)\neq 0$ and

$$|e^n(k|k)(2+\psi(k)^2x_r(k-d)^t x_r(k-d))|\geq|2\Delta^n(k)| \tag{6.43}$$

Condition (6.43) may be rewritten in the form

$$|e^n(k|k)|\geq|2\Delta^n(k)|/(2+\psi(k)^2x_r(k-d)^t x_r(k-d)) \tag{6.44}$$

Using (6.36), inequality (6.44) may be expressed as

$$|e^n(k|k-1)|\geq\frac{2+2\psi(k)^2x_r(k-d)^t x_r(k-d)}{2+\psi(k)^2x_r(k-d)^t x_r(k-d)}\cdot|\Delta^n(k)| \tag{6.45}$$

It may be shown that the following inequality is true for all $\psi(k)\neq 0$ that verify conditions (6.30)-(6.33) of the normalized estimation system

$$|e^n(k|k-1)|\geq\frac{2+2\psi(k)^2x_r(k-d)^t x_r(k-d)}{2+\psi(k)^2x_r(k-d)^t x_r(k-d)}\cdot\Delta_b \tag{6.46}$$

where according to (6.29) $\Delta_b>|\Delta^n(k)|\ \forall k$.

Thus from (6.29) and (6.46), condition (6.45) is satisfied if $\psi(k)\neq 0$. This completes the proof of Lemma 6.2.

Lemma 6.3: Along the solution of the adaptive algorithm (6.25)-(6.33), $e^n(k|k)$ verifies

(i) $|e^n(k|k)|\leq\Delta_b'(\psi_\ell,\Delta_b,k)<2\Delta_b$ for those k for which $\psi(k)=0$.

(ii) If the number of sampling times for which $\psi(k)\neq 0$ tends to infinity as $k\to\infty$ then $\lim\limits_{k\to\infty} e^n(k|k)=0$ for those k for which $\psi(k)\neq 0$.

Proof: From (6.36), $e^n(k|k)=e^n(k|k-1)$ when $\psi(k)=0$. Then using condition (6.27), property (i) of this lemma is easily deduced.

Consider that the number of sampling times for which $\psi(k)\neq 0$ tends to infinity as $k\to\infty$. Since $||\tilde{\theta}_r(k)||^2$ is a bounded nonincreasing function it converges and therefore

$$\lim_{k\to\infty}[||\tilde{\theta}_r(k)||^2-||\tilde{\theta}_r(k-1)||^2]=0 \qquad (6.47)$$

Using (6.42) and (6.47) gives

$$\lim_{k\to\infty}[-\psi(k)^2e^n(k|k)^2(2+\psi(k)^2x_r(k-d)^tx_r(k-d)) +2\psi(k)^2e^n(k|k)\Delta^n(k)]=0 \qquad (6.48)$$

(6.48) implies that for those k for which $\psi(k)\neq 0$

$$e^n(k|k)^2(2+\psi(k)^2x_r(k-d)^tx_r(k-d))\to 2e^n(k|k)\Delta^n(k) \qquad (6.49)$$

as $k\to\infty$

From (6.36) and (6.46) it may be deduced that

$$|e^n(k|k)|(2+\psi(k)^2x_r(k-d)^tx_r(k-d))\geq 2\Delta_b \qquad (6.50)$$

for all $\psi(k)\neq 0$.

Since by definition $\Delta_b>\Delta_m$ in (6.29), there exists $\delta>0$ such that

$$\Delta_b-\Delta_m=\delta \tag{6.51}$$

Using (6.50) and (6.51) it follows that

$$|e^n(k|k)|(2+\psi(k)^2x_r(k-d)^tx_r(k-d))\geq 2\delta + 2|\Delta^n(k)| \tag{6.52}$$

for all $\psi(k)\neq 0$

Since $\delta>0$, (6.52) proves that, for those k for which $\psi(k)\neq 0$, $e^n(k|k)(2+\psi(k)^2x_r(k-d)^tx_r(k-d))$ cannot tend to $2\Delta^n(k)$ as $k\to\infty$, and therefore (6.49) implies that $\lim_{k\to\infty} e^n(k|k)=0$ for those k for which $\psi(k)\neq 0$, which completes the proof of this lemma.

Lemma 6.4: Along the solution of the adaptive algorithm (6.25)-(6.33):

$$\lim_{k\to\infty} [\hat{\theta}_r(k)-\hat{\theta}_r(k-h)]=0$$

where h is a finite integer.

Proof: Using (6.41) and (6.47) gives

$$\lim_{k\to\infty} [2\psi(k)^2e^n(k|k)(e^n(k|k)-\Delta^n(k))+ \psi(k)^4e^n(k|k)^2x_r(k-d)^tx_r(k-d)]=0 \tag{6.53}$$

From the parameter estimation law (6.37)

$$||\hat{\theta}_r(k)-\hat{\theta}_r(k-1)||^2=\psi(k)^4e^n(k|k)^2x_r(k-d)^tx_r(k-d) \tag{6.54}$$

Now using (6.53) and (6.54)

$$\lim_{k\to\infty} [2\psi(k)^2e^n(k|k)(e^n(k|k)-\Delta^n(k))+||\hat{\theta}_r(k)-\hat{\theta}_r(k-1)||^2]=0 \tag{6.55}$$

Since $\{|\Delta^n(k)|\}$ is bounded from Lemma 6.1, using properties (i) and (ii) of Lemma 6.3, it may be deduced that

$$\lim_{k\to\infty} [2\psi(k)^2e^n(k|k)(e^n(k|k)-\Delta^n(k))]=0 \tag{6.56}$$

and combining (6.55) and (6.56) gives

$$\lim_{k\to\infty} ||\hat{\theta}_r(k)-\hat{\theta}_r(k-1)||=0 \tag{6.57}$$

Since

$$||\hat{\theta}_r(k)-\hat{\theta}_r(k-h)||\le\sum_{j=1}^{h}||\hat{\theta}_r(k+1-j)-\hat{\theta}_r(k-j)|| \tag{6.58}$$

Taking limits on both sides of (6.58) as k→∞ and using (6.57) yields

$$\lim_{k\to\infty} ||\hat{\theta}_r(k)-\hat{\theta}_r(k-h)||=0$$

which proves this lemma.

6.3.3 Properties of the Unnormalized a posteriori Estimation Error

Lemma 6.5: The unnormalized *a posteriori* estimation error $e(k|k)$ defined by equation (6.7), under the adaptation law defined by (6.25)-(6.33) verifies the following:

there exists $k_2<\infty$ such that

$$|e(k|k)|<2\Delta_b\cdot\max(||\Phi(k-d)||,c) \tag{6.59}$$

$\forall k$ such that $0<k_2<k<\infty$

Proof: From (6.7) and (6.34) we may write

$$e^n(k|k)=e(k|k)/n(k) \tag{6.60}$$

It may be concluded from Lemma 6.3 that there exists k_2 such that

$$|e^n(k|k)|<2\Delta_b \qquad \forall k>k_2 \tag{6.61}$$

Then from (6.60) and (6.61)

$$|e(k|k)|<2\Delta_b\cdot n(k) \tag{6.62}$$

and from the definition of n(k) in (6.19)

$$|e(k|k)|<2\Delta_b\cdot\max(||\Phi(k-d)||,c) \qquad \forall k>k_2 \tag{6.63}$$

which completes the proof of this lemma.

6.4 Global Stability Result

The global stablity result of the adaptive predictive control system in the presence of unmodeled dynamics and bounded noise and disturbances is contained in Theorem 6.2. First the boundedness of the I/O vector $\Phi(k)$ is proven by showing that, from the results of Lemma 6.4 and 6.5, the condition of Theorem 6.1 is satisfied. Then it is shown that the estimated parameters converge to a set of constant values and that the absolute value of the tracking error is bounded.

Theorem 6.2: The adaptive predictive control system described by the adaptation mechanism defined in (6.25)-(6.33) and the control law defined by (6.8) applied to the process described by (6.18) and verifying the following condition:

$$a_1 > 2\Delta_b \qquad (6.64)$$

has the following properties:

a) $||\Phi(k)||<\infty \quad \forall k>0$

b) there exists $k_f<\infty$ such that $\hat{\theta}_r(k)=\hat{\theta}_r(k-1) \quad k>k_f$

c) $|\epsilon(k)|=|e(k|k)|<2\Delta_b\cdot\max(||\Phi(k-d)||,c) \quad \forall k>k_f+d-1$

Proof: From Lemma 6.5

$$\frac{|e(k|k)|}{\max(||\Phi(k-d)||,c)} < 2\Delta_b \qquad (6.65)$$

Then from condition (6.64)

$$a_1 > 2\Delta_b > \frac{|e(k|k)|}{\max(||\Phi(k-d)||,c)} \qquad (6.66)$$

Since $\lim_{k\to\infty} ||\hat{\theta}_r(k)-\hat{\theta}_r(k-d)||=0$ according to Lemma 6.4, (6.66) implies that the condition of Theorem 6.1 is satisfied for some finite k_1, which proves property (a).

Since $\{||x_r(k)||\}$ and $\{\psi(k)^2\}$ are bounded, if the number of sampling times for which $\psi(k)\neq 0$ tends to infinity as $k\to\infty$, from property (ii) of Lemma 6.3 it may deduced that

$$\lim_{k\to\infty} [|e^n(k|k)|(2+\psi(k)^2 x_r(k-d)^t x_r(k-d))]=0 \qquad (6.67)$$

for those k for which $\psi(k)\neq 0$.

However (6.67) contradicts (6.50) where $\Delta_b>0$. Therefore the number of times k for which $\psi(k)\neq 0$ cannot tend to infinity, i.e. there exists a finite k_f such that $\forall k>k_f$, $\psi(k)=0$. Using (6.26) gives

$$\hat{\theta}_r(k)=\hat{\theta}_r(k-1) \qquad \forall k>k_f \qquad (6.68)$$

which proves property (b). Using this property

$$\hat{\theta}_r(k)=\hat{\theta}_r(k-d) \qquad \forall k>k_f+d-1 \qquad (6.69)$$

Now using (6.11) and (6.69)

$$\epsilon(k)=e(k|k) \qquad \forall k>k_f+d-1 \qquad (6.70)$$

and combining (6.70) with the result of Lemma 6.5 completes the proof of this theorem.

6.5 Simulation Results

The adaptive control system developed in this chapter is applied to a benchmark example which has been used extensively throughout the literature (e.g. Astrom, 1984; Johnson *et al*, 1984; Kosut and Friedlander, 1985) to analyze the robustness of various adaptive controllers. This example consists of a third order plant which includes second order, high frequency dynamics

$$y(s) = \frac{2}{s+1} \cdot \frac{229}{s^2+30s+229} \cdot u(s) \qquad (6.71)$$

This plant was analyzed by Rohrs *et al* (1984) at two different sampling rates (T_s=.04s,.4s) and further analysis of this example was performed by Cluett *et al* (1986) over a range of sampling times. A plot of the zero locations of this plant as a function of sampling time in Cluett *et al* (1986) showed that an unstable zero appears in the discrete-time plant representation of (6.71) for T_s<0.2s. (The equivalent discrete-time systems were obtained by the standard technique called *hold equivalence.*) Therefore in order to satisfy the stable-inverse plant assumption (subsection 6.2.3, assumption 3), the system must be analyzed at sampling periods greater than 0.2s. Two different sampling periods of .3s and .25s were used. For T_s=.3s, the discrete-time representation of the plant is

$$\begin{aligned} y(k) = {} & .759y(k-1)-.0137y(k-2)+.0000914y(k-3) \\ & +.313u(k-1)+.193u(k-2)+.00292u(k-3) \qquad (6.72) \end{aligned}$$

and for T_s=.25s,

$$\begin{aligned} y(k) = {} & .820y(k-1)-.0327y(k-2)+.000431y(k-3) \\ & +.234u(k-1)+.185u(k-2)+.00479u(k-3) \qquad (6.73) \end{aligned}$$

In one set of experiments (Figures 6.1 and 6.2) the new APCS algorithm (A1) presented in this chapter, that includes the normalized estimation system with the on/off parameter adaptation feature, is applied to the plant in (6.71) under the discrete-time representations (6.72) (Figure 6.1) and

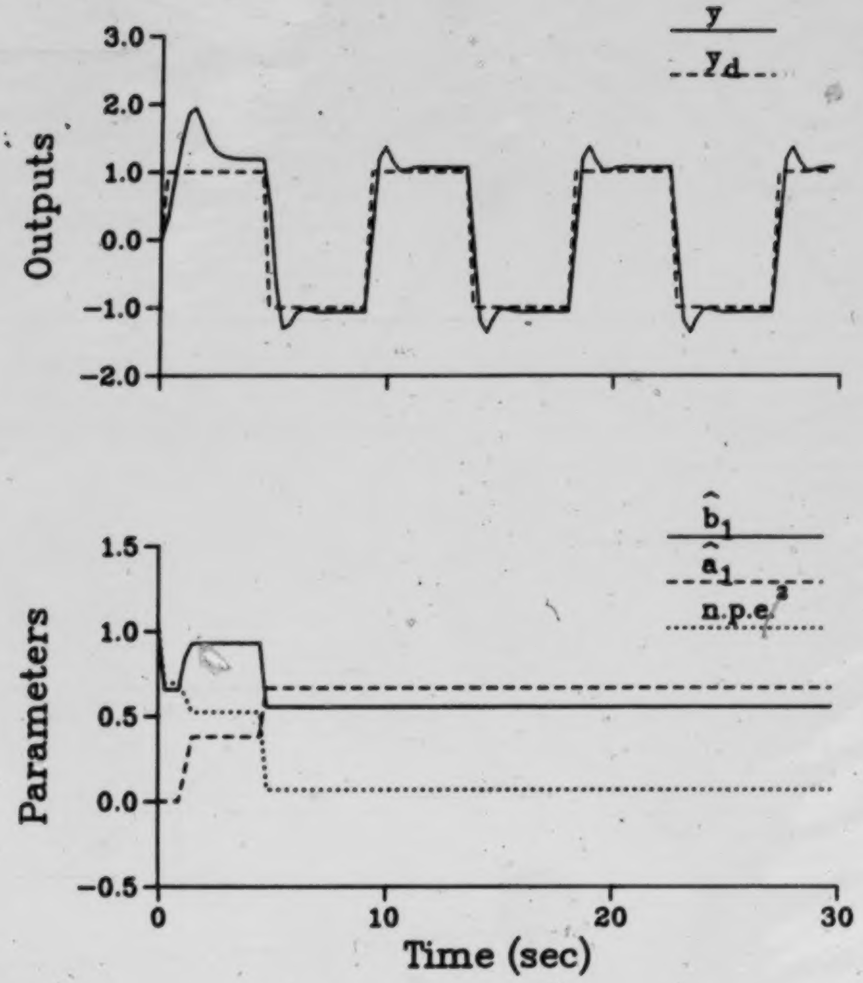

Figure 6.1 Simulation of A1 with T_S=0.3s

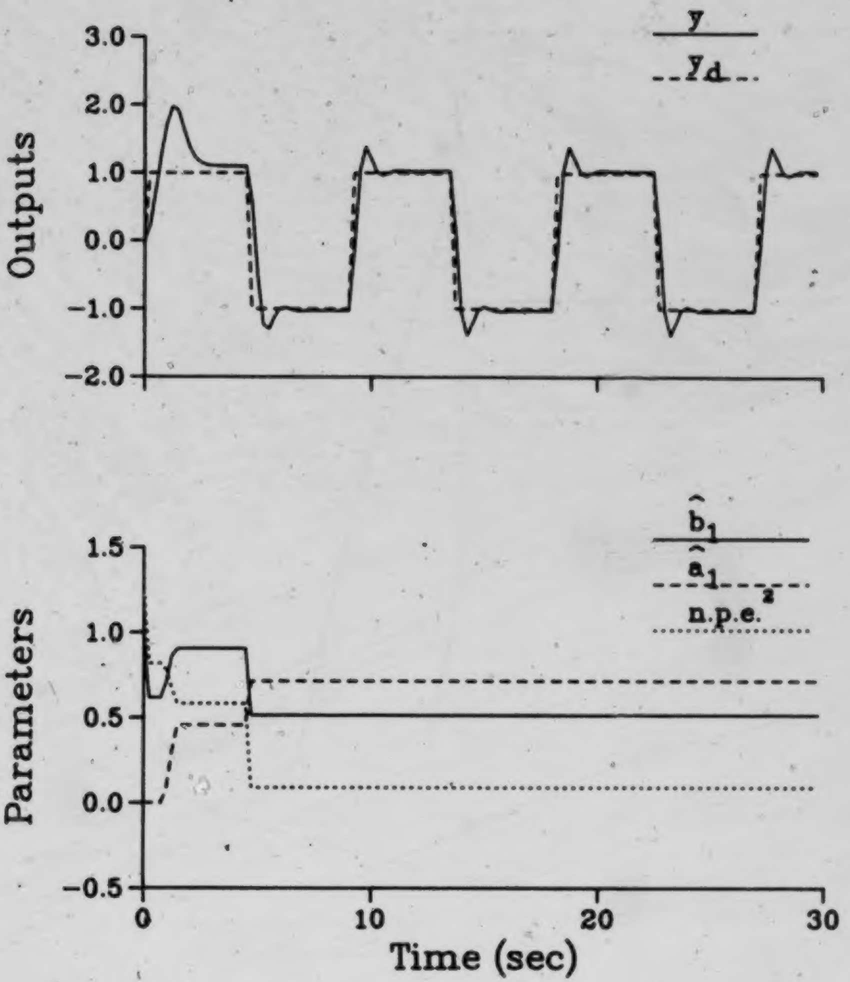

Figure 6.2 Simulation of A1 with T_s=0.25s

(6.73) (Figure 6.2). The adaptive predictive model was assumed to be first order ($n_r=2$). As may be seen from equations (6.72) and (6.73), the actual process is third order ($n_p=6$). This mismatch between the model order and the actual process order results in the presence of unmodeled dynamics. The first order plant assumption produces a control law of the form

$$y_d(k+1) = \hat{\theta}_r(k)^t \phi_r(k) \tag{6.74}$$

where

$$\hat{\theta}_r(k)^t = [\hat{a}_1(k) \quad \hat{b}_1(k)]$$

$$\phi_r(k)^t = [y(k) \quad u(k)]$$

The scalar function $\psi(k)^2$ defined in subsection 6.3.1 was always chosen to be equal to $\psi_b(k)^2$ as defined in equations (6.31) and (6.33), with $\psi_\ell{}^2=0.1$, $\psi_u{}^2=1.0$ and the upper bound Δ_b evaluated using the expression on the righthand side of (6.24). The parameter c required in the normalization (6.19) was chosen equal to 1. The parameter m, which must be greater than or equal to n_p, was selected to be equal to 8.

For comparison purposes, a second set of experiments (Figures 6.3 and 6.4) was performed using the APCS originally presented by Martin-Sanchez (1976a) (A2) (which corresponds to the deterministic case presented by Martin-Sanchez (1982) and Martin-Sanchez *et al* (1984) with $\psi(k)^2=1.0\ \forall k$) applied to the plant in (6.71) under the discrete-time representations (6.72) (Figure 6.3) and (6.73) (Figure 6.4). This type of 'deterministic' control algorithm has been the subject of analysis in previous robustness

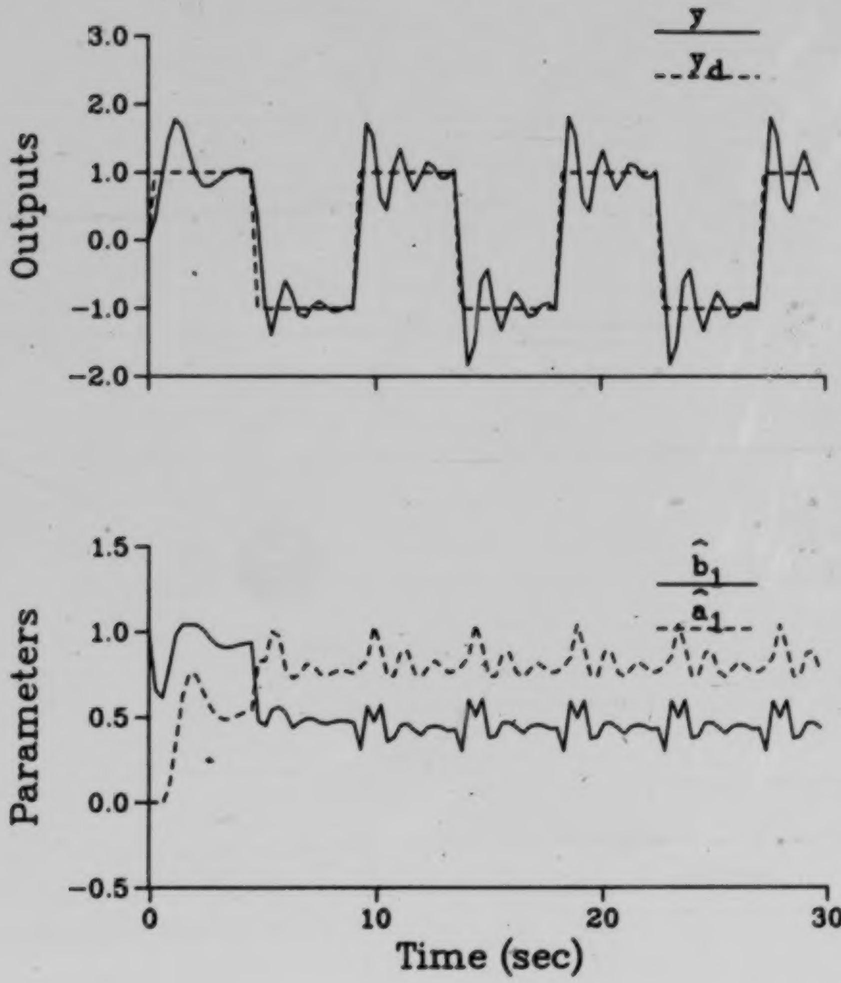

Figure 6.3 Simulation of A2 with T_s=0.3s

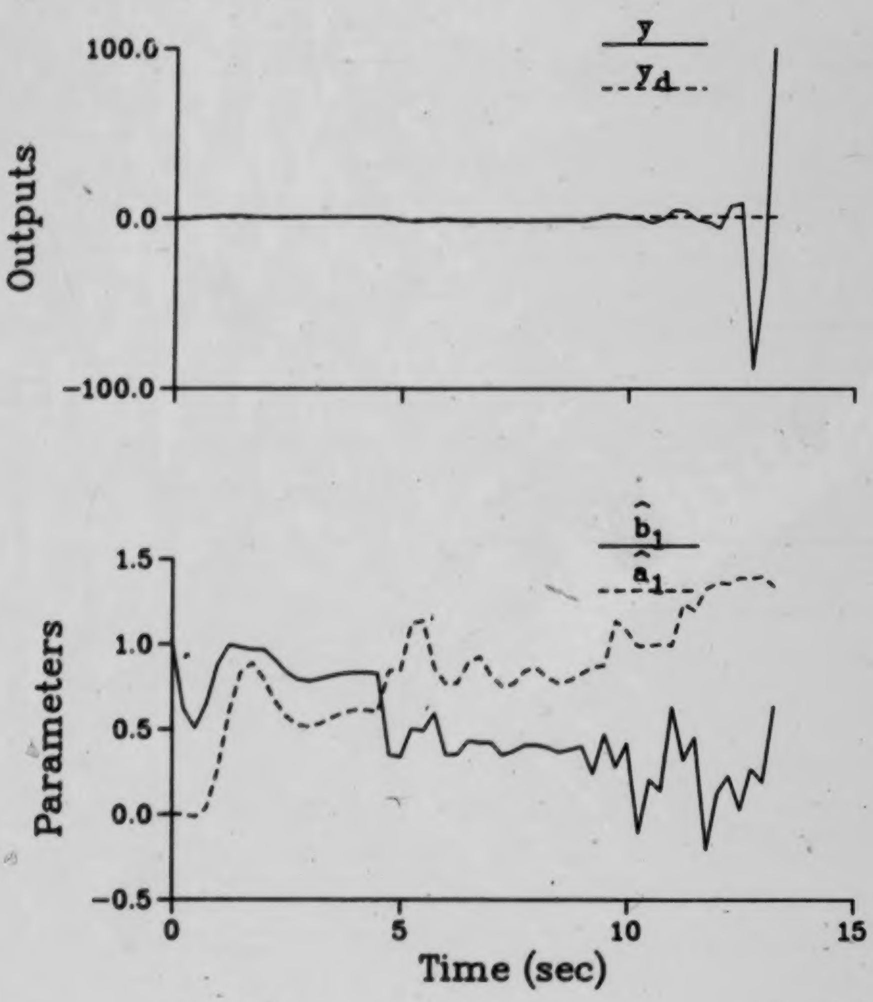

Figure 6.4 Simulation of A2 with T_s=0.25s

studies where stability problems due to unmodeled dynamics were demonstrated. Therefore the APCS algorithm used for this set of experiments is given by

$$\begin{aligned} \hat{\theta}_r(k) &= \hat{\theta}_r(k-1)+\frac{[y(k)-\hat{\theta}_r(k-1)^t\phi_r(k-1)]\cdot\phi_r(k-1)}{1+\phi_r(k-1)^t\phi_r(k-1)} \\ y_d(k+1) &= \hat{\theta}_r(k)^t\phi_r(k) \end{aligned} \qquad \text{(A2)}$$

The AP model was also assumed to be first order in this case. Consequently the control law in (A2) is equivalent to (6.74) but in this case the normalized estimation scheme with the on/off criterion was not used. In all the experiments the same square wave signal for y_d was used. However note that the scale in Figure 6.4 has been changed to better illustrate the results.

6.6 Discussion of Results

Figure 6.3 shows the simulation results of applying A2 for T_s=0.3s. The system is stable for this case but the response is oscillatory. Figure 6.4 is the result of applying A2 with the faster sampling rate (T_s=0.25s). In this experiment, the system becomes unstable. As illustrated by these results, larger sampling periods generally lead to a more robust control system. On the other hand, even a small decrease in the sampling period may initiate unstable behaviour.

Figures 6.1 and 6.2 show the results of applying A1 with sampling periods of 0.3s and 0.25s respectively. The results in Figures 6.1 and 6.2 demonstrate that the three properties defined by Theorem 6,2 are obtained: 1) the system is stable (i.e. the I/O vector is bounded), 2) the parameters converge in a finite time to a tuned set, and 3) the norm of the tracking error is bounded. The plots of the adaptive predictive model parameters also include the squared norm of the parameter error (n.p.e.^{2}) which was defined previously in Lemma 6.2. As proven in this lemma, the squared norm is a nonincreasing sequence along the solution of the normalized estimation system. This property is clearly demonstrated by the n.p.e.^{2} plots in Figures 6.1 and 6.2. These experiments show how A1 overcomes the destabilizing effects of unmodeled dynamics which appear in the A2 results. Furthermore, the behaviour of the control and estimation system under A1 illustrate the theoretical results presented in this chapter.

The effect of the choice of the AP model order on the performance of the control system becomes clear by analyzing expression (6.24) in connection with (6.22), and property (c) of Theorem 6.2. In fact (6.24) shows how a large difference between the actual process order (n_p) and the reduced model order (n_r) demands a greater upper bound Δ_b on the normalized perturbation sequence $\{\Delta^n(k)\}$. This implies, by property (c) of Theorem 6.2, that the asymptotic convergence region for the tracking error becomes larger. On the other hand, if the order mismatch (n_p-n_r) is smaller, then the norm of the unmodeled parameter vector, $||\theta_u||$, is also smaller and the required upper bound Δ_b may be decreased. The effect of decreasing Δ_b reduces the upper bound on the tracking error norm according to property (c) and hence improves the performance of the system. In the matched, deterministic case ($\Delta_b=0$) the tracking error asymptotically converges to zero. This illustrates the direct link which exists between decreased unmodeled dynamics and improved controller performance.

The parameter α_1, first introduced in assumption 3, is an inherent parameter in the dynamics of each stable-inverse process. The theoretical results presented here relate this inherent property of the process to the stability result under a reduced adaptive predictive control structure. The practical interpretation reduces to the fact that if α_1 is large, then there is a large margin for a reduced control structure that will result in a stable system. If the process has a small α_1, it may happen that even a small model reduction may not be tolerated. The theoretical results presented in this chapter, particularly condition (6.64) in Theorem 6.2, show that the inherent nature of the process may limit the allowable model reduction. The experimental results obtained using a benchmark example have illustrated that, under these limits, the stability problem related to unmodeled dynamics is overcome.

Condition (6.64) may be interpreted as a "sufficiently small" requirement for the unmodeled dynamics. A very similar condition is used by Goodwin *et al* (1985a) in their recent global stability result for an adaptive control algorithm applied to a class of plants with unmodeled dynamics.

The influence of the sampling rate on key parameters such as α_1 and $||\theta_u||$, as well as the effect of an incremental implementation of APCS on the elimination of the bias in the tracking error, may be the subject of future studies. However, in practical applications where the process is unknown and/or technical conditions of Theorems 6.1 and 6.2 (such as linearity of the real process) are not met then a value for Δ_b is most reasonably obtained using its physical interpretation as an upper bound on $\Delta(k)$.

6.7 A Robust Adaptive Controller

Some of the earliest work on the issue of adaptive controller robustness to unmodeled dynamics was done by Lim (1982), and Gawthrop and Lim (1982) on the self-tuning controller (Clarke and Gawthrop, 1975). Conditions were derived for stability of the adaptive system in terms of the design parameters available in the STC formulation. The work of Lim (1982) contains one major drawback. The stability results derived in these papers require an *a priori* plant signal boundedness assumption. This defect was removed by Ortega *et al* (1985) by including normalized signals in the least squares parameter adaptation algorithm. However the control law in Ortega *et al* (1985) lacks the original STC weighting polynomials, P, Q and R.

In this chapter a general approach for the stability analysis of adaptive predictive control systems has been presented based on a theorem which states a *single* condition for stability in terms of convergence properties of the estimation error and the estimated parameters. This approach was used to prove global stability of an adaptive controller in the presence of bounded disturbances and unmodeled dynamics due to process-model order mismatch. A normalized estimation system, based on a gradient scheme, with an on/off parameter adaptation criterion was shown to verify the convergence properties that guarantee global stability.

Much of the convergence analysis to date dealing with adaptive control systems in the presence of unmodeled dynamics has been applied only to the model reference-type of control law using gradient parameter estimators. Goodwin *et al* (1984) pointed out that the model reference controller is fundamentally non-robust and that the gradient procedure is known to converge slowly. From the practitioners viewpoint, more work needs to be done on the analysis of more widely used adaptive controllers such as STC.

This section extends the stability results presented earlier in this chapter to include; 1) a normalized least squares parameter estimation scheme with a parameter adaptation stopping criterion, and 2) the concept of an augmented plant which enables P, Q and R weighting polynomials to be incorporated into the predictive control law. The end result of these extensions is a STC-type of adaptive controller which is globally stable in the presence of unmodeled dynamics due to process-model order mismatch.

6.7.1 A Normalized Least Squares Parameter Estimator

Consider the following least squares parameter estimation scheme in place of the gradient scheme in (6.26). The estimated parameter vector is updated according to

$$\hat{\theta}_r(k) = \hat{\theta}_r(k-1) + \frac{\psi(k)^2 P(k-1) x_r(k-d) e^n(k|k-1)}{1+\psi(k)^2 x_r(k-d)^t P(k-1) x_r(k-d)} \qquad (6.75)$$

and the time-varying matrix P(k) is updated by

$$P(k) = P(k-1) - \frac{\psi(k)^2 P(k-1) x_r(k-d) x_r(k-d)^t P(k-1)}{1+\psi(k)^2 x_r(k-d)^t P(k-1) x_r(k-d)} \qquad (6.76)$$

The scalar $\psi(k)^2$ is determined by a criterion for stopping or continuing parameter adaptation.

(a) $\psi(k)^2=1$ if

$|e^n(k|k-1)|>(1+x_r(k-d)^t P(k-1) x_r(k-d))^{1/2}\cdot\Delta_b$

Δ_b is an estimate of an upper bound on the absolute value of the perturbation signal $\Delta^n(k)$ (i.e. $\Delta_b>|\Delta^n(k)|\ \forall k$).

(b) $\psi(k)^2=0$ if

$|e^n(k|k-1)|\leq(1+x_r(k-d)^t P(k-1) x_r(k-d))^{1/2}\cdot\Delta_b$

The adaptive algorithm presented in (6.75) and (6.76) is a normalized version of a least squares parameter estimation scheme presented by Goodwin and Sin (1984). The nonzero value for $\psi(k)^2$ could be selected to be any positive constant or could be chosen from a range of positive quantities as in subsection 6.3.1. For clarity a value of unity has been used here.

6.7.2 Convergence Properties of the Normalized System

The *a posteriori* estimation error for the normalized system in (6.22) is defined as

$$e^n(k|k) = y^n(k)-\hat{\theta}_r(k)^t x_r(k-d) \qquad (6.77)$$

Subtracting (6.77) from (6.25) gives

$$e^n(k|k-1) = [\hat{\theta}_r(k)-\hat{\theta}_r(k-1)]^t x_r(k-d)+e^n(k|k) \qquad (6.78)$$

Substituting (6.75) into (6.78) yields

$$e^n(k|k)=e^n(k|k-1)/(1+\psi(k)^2 x_r(k-d)^t P(k-1)x_r(k-d)) \qquad (6.79)$$

and the parameter update equation (6.75) may then be written as

$$\hat{\theta}_r(k) = \hat{\theta}_r(k-1)+\psi(k)^2 P(k-1)x_r(k-d)e^n(k|k) \qquad (6.80)$$

Lemma 6.6: Along the solution of the adaptive algorithm (6.75) and (6.76):

(a) $V(k)-V(k-1) \le 0 \quad \forall k$

where $V(k)=\tilde{\theta}_r(k)^t P(k)^{-1}\tilde{\theta}_r(k)$ and $\tilde{\theta}_r(k)=\theta_r-\hat{\theta}_r(k)$.

(b) The number of times k for which $\psi(k)^2 \neq 0$ cannot tend to infinity, i.e. there exists a finite k_f such that $\forall k>k_f$, $\psi(k)^2=0$.

Proof:

(a) Subtracting θ_r from both sides of (6.80) gives

$$\tilde{\theta}_r(k) = \tilde{\theta}_r(k-1)-\psi(k)^2 P(k-1)x_r(k-d)e^n(k|k) \qquad (6.81)$$

Multiplying through by $\tilde{\theta}_r(k)^t P(k-1)^{-1}$ yields

$$\tilde{\theta}_r(k)^t P(k-1)^{-1}\tilde{\theta}_r(k)=V(k-1)-2\psi(k)^2\tilde{\theta}_r(k)^t x_r(k-d)e^n(k|k) \\ -\psi(k)^4 x_r(k-d)^t P(k-1)x_r(k-d)e^n(k|k)^2 \qquad (6.82)$$

From (6.76) and the Matrix Inversion Lemma

$$V(k)-V(k-1) = \psi(k)^2[\tilde{\theta}_r(k)^t x_r(k-d)]^2 \\ -2\psi(k)^2\tilde{\theta}_r(k)^t x_r(k-d)e^n(k|k) \\ -\psi(k)^4 x_r(k-d)^t P(k-1)x_r(k-d)e^n(k|k)^2 \quad (6.83)$$

Using (6.22) and (6.77), $e^n(k|k)$ may be expressed as

$$e^n(k|k) = \tilde{\theta}_r(k)^t x_r(k-d) + \Delta^n(k) \quad (6.84)$$

Combining (6.83) and (6.84) gives

$$V(k)-V(k-1)=\psi(k)^2\Delta^n(k)^2 \quad (6.85)$$
$$-\psi(k)^2 e^n(k|k)^2[1+\psi(k)^2 x_r(k-d)^t P(k-1)x_r(k-d)]$$

The left hand side of (6.85) is equal to zero if $\psi(k)^2=0$ and less than zero if $\psi(k)^2=1$ and

$$e^n(k|k)^2[1+x_r(k-d)^t P(k-1)x_r(k-d)] > \Delta^n(k)^2 \quad (6.86)$$

The following inequality is true for all $\psi(k)^2=1$

$$|e^n(k|k-1)|>[1+x_r(k-d)^t P(k-1)x_r(k-d)]^{1/2}\cdot\Delta_b \quad (6.87)$$

Using (6.79), inequality (6.87) may be expressed as

$$e^n(k|k)^2[1+x_r(k-d)^t P(k-1)x_r(k-d)]>\Delta_b^2 \quad (6.88)$$

Since by definition $\Delta_b>|\Delta^n(k)|\ \forall k$ condition (6.86) is satisfied if $\psi(k)^2=1$. This completes the proof of part (a).

(b) Consider that the number of sampling times for which $\psi(k)^2\neq 0$ tends to infinity as $k\to\infty$. Since $V(k)$ is a bounded nonincreasing function, it converges and therefore

$$\lim_{k\to\infty}[V(k)-V(k-1)] = 0 \quad (6.89)$$

Using (6.85) and (6.89) implies that

$$e^n(k|k)^2[1+x_r(k-d)^t P(k-1)x_r(k-d)]\to\Delta^n(k)^2 \quad (6.90)$$

as $k\to\infty$ for those k for which $\psi(k)^2\neq 0$

From (6.88) it may be concluded that

$$e^n(k|k)^2[1+x_r(k-d)^t P(k-1)x_r(k-d)]>\Delta_b^2 \quad (6.91)$$

for all $\psi(k)^2 \neq 0$

However (6.91) contradicts (6.90) where $\Delta_b^2 > \Delta^n(k)^2$. Therefore the number of times k for which $\psi(k)^2 \neq 0$ cannot tend to infinity, i.e. there exists a finite k_f such that $\forall k > k_f$, $\psi(k)^2 = 0$. This completes the proof of Lemma 6.6.

6.7.3 Global Stability Result

Theorem 6.3: The adaptive control system described by the adaptive mechanism defined in (6.75) and (6.76) and the control law defined by (6.8) applied to the process described in (6.18) is globally stable (i.e. $||\Phi(k)|| < \infty \quad \forall k$) if the following condition is verified:

$$a_1 > [1 + n_r \cdot \lambda_{max}]^{1/2} \cdot \Delta_b \tag{6.92}$$

where λ_{max} is the maximum eigenvalue of P(0).

Proof: From (6.79) and the result of Lemma 6.6(b),

$$|e^n(k|k)| \leq [1 + x_r(k-d)^t P(k-1) x_r(k-d)]^{1/2} \cdot \Delta_b \tag{6.93}$$

$\forall k > k_f$

From (6.60)

$$e^n(k|k) = e(k|k)/n(k) \tag{6.94}$$

Combining (6.93) and (6.94) gives

$$|e(k|k)| \leq [1 + x_r(k-d)^t P(k-1) x_r(k-d)]^{1/2} \cdot \Delta_b \cdot n(k) \tag{6.95}$$

$\forall k > k_f$

From the definition of n(k) in (6.19)

$$n(k) \leq \max(||\Phi(k-d)||, c) \tag{6.96}$$

and from the definition of $x_r(k-d)$ in (6.21) and of P(k) in (6.76) it is clear that

$$x_r(k-d)^t P(k-1) x_r(k-d) \le \lambda_{max} \cdot ||x_r(k-d)||^2$$
$$\le \lambda_{max} \cdot n_r \qquad (6.97)$$

From (6.95)-(6.97) and the condition in (6.92), it may be concluded that

$$\alpha_1 > \frac{|e(k|k)|}{\max(||\Phi(k-d)||, c)} \qquad (6.98)$$

$$\forall k > k_f$$

Since the number of times k for which $\psi(k)^2 \neq 0$ cannot tend to infinity from Lemma 6.6(b), then there exists some finite $k_1 = k_f + d - 1$ such that

$$\hat{\theta}_r(k) = \hat{\theta}_r(k-d) \quad \forall k > k_1 \qquad (6.99)$$

Therefore the stability condition of Theorem 6.1 is satisfied $\forall k > k_1$ which completes the proof of this theorem.

The condition in (6.92) is similar to the condition in (6.64). It is interesting to note that (6.92) depends on the initial choice by the user of the matrix P. A small value for λ_{max} is desirable from a robustness point of view but would result in slower parameter convergence and hence a tradeoff is necessary.

It is well known that the least squares estimation scheme in (6.75) and (6.76) must be modified for practical use. For instance, P(k) must be prevented from going to zero in order that the algorithm be able to track process parameter variations and must also be prevented from "blowing up". Two methods which overcome this difficulty and also retain the important convergence properties of Lemma 6.6 are: (i) the constant trace algorithm of Sripada and

Fisher (1986) which uses a variable forgetting factor to control the trace, and (ii) the regularized constant trace algorithm of Goodwin *et al* (1985b) which includes features of both gradient and least squares.

6.7.4 A Self-Tuning Control Law

One method of including some penalty or weighting on the output, y, and the input, u, in the control system framework presented in this chapter is to define an *augmented* plant with output, z, and input, u, where

$$z(k) = P(q^{-1})y(k) + q^{-d}Q(q^{-1})u(k) \tag{6.100}$$

and P and Q are polynomials in the delay operator q^{-1}. (P and Q may also be considered as transfer functions (Clarke and Gawthrop, 1979) but for clarity will be assumed to be polynomials.) The process representation in (6.18) may be rewritten as

$$A(q^{-1})y(k) = q^{-d}B(q^{-1})u(k) + \xi(k) \tag{6.101}$$

where A and B are polynomials in the delay operator and ξ represents the effect of disturbances and noise. Substituting (6.101) into (6.100) yields

$$Az(k) = q^{-d}(PB + QA)u(k) + P\xi(k) \tag{6.102}$$

z(k) may now be expressed in a form similar to (6.18), i.e.

$$z(k) = \theta_z{}^t\phi_z(k-d) + \Delta_z(k) \tag{6.103}$$

where

$$\phi_z(k-d)^t = [z(k-d)\ z(k-d-1)\cdots u(k-d)\ u(k-d-1)\cdots]$$

θ_z contains the parameters of the augmented process and $\Delta_z(k)$ represents the effect of disturbances and noise on the output, z, at time k. The adaptive predictive control system in subsection 6.2.2 may now be applied to the augmented plant with $y(k)$ replaced throughout by $z(k)$. The predictive control law itself becomes

$$y_{df}(k+d) = \hat{\theta}_{zr}(k)^t\phi_{zr}(k) \tag{6.104}$$

where

$$\phi_{zr}(k)^t = [z(k)\ z(k-1)\cdots u(k)\ u(k-1)\cdots]$$

and the dimension of ϕ_{zr} and $\hat{\theta}_{zr}$ may be less than or equal to the dimension of ϕ_z and θ_z in (6.103). y_{df} represents a filtered setpoint value equal to $R(q^{-1})y_d$ where R is a polynomial in the delay operator. It is worth noting that the original control system presented in subsection 6.2.2 may be viewed as a subset of this more general augmented plant approach by setting $P=1$, $Q=0$ and $R=1$.

The choice of z as the output of an augmented system is meaningful from a *qualitative* point of view because minimum variance control of the augmented plant (i.e. $J=E\{z-Ry_d\}^2$) implies minimization of an index $J=E\{Py+Qu-Ry_d\}^2$ for the original system. This latter index was first suggested by Clarke and Gawthrop (1975) for self-tuning controllers. The adaptive controller presented here is however not minimum variance because of the use of the parameter adaptation stopping criterion.

The role of Q has two benefits: 1) Q may be used to reduce the excessive manipulation of the control input, $u(k)$, which often results from predictive control when there is no weighting of the control action, and 2) some nonminimum phase processes can be handled since the zeros of the augmented system (i.e. the roots of PB+QA) can be shifted into the unit circle if Q is selected correctly and the original plant is open-loop stable. This interpretation of zero shifting using an augmented plant was first presented by Johnstone *et al* (1980). An alternative augmented plant representation is the generalized output considered by Clarke and Gawthrop (1975), and Clarke (1984).

The tracking error for the augmented control system is defined as

$$\begin{aligned} e(k) &= z(k) - y_{df}(k) \qquad (6.105)\\ &= Py(k)+q^{-d}Qu(k)-Ry_d(k) \end{aligned}$$

If $Q=\lambda(1-q^{-1})$, where λ is a scalar, which puts weighting on *changes* in control action (Clarke, 1984), and if the input has reached a steady value, then

$$e(k) \simeq Py(k)-Ry_d(k) \qquad (6.106)$$

If the tracking error, under the adaptive control scheme, converges to zero, then

$$y(k) \to \frac{R}{P} \cdot y_d(k) \qquad (6.107)$$

and the interpretation of R/P as a reference model for the closed-loop response is obvious.

6.8 Conclusions

1. The single condition for stability of adaptive predictive control systems (APCS) presented in Theorem 6.1 encompasses all previous APCS stability results and is applicable to a

broad class of deterministic, stochastic and time-varying linear systems.

2. Stability of APCS in the presence of process-model order mismatch has been proven using a normalized parameter estimation scheme, which permits a formal proof that the modeling errors can be treated as a bounded disturbance, and a parameter adaptation stopping criterion.

3. The above theoretical results confirm the intuitive expectation that the degree of process-model order mismatch that can be introduced into an APCS without violating the guarantee of global stability is limited by inherent properties of the process. These process properties (e.g. a_1 of assumption 3) deserve further investigation.

4. The assumption of a tuned set of controller parameters and/or the strictly positive real or conicity condition used by other authors is not required in this approach. However, Theorem 6.2 shows that parameter convergence to such a tuned set is a natural result of the approach used herein.

5. Simulation results using a benchmark example illustrate the theoretical properties, practicality and performance of the proposed approach.

6. The above stability results have been extended to a

self-tuning controller (STC) type adaptive system with a least squares parameter estimation scheme. This system uses an STC-type control law with P, Q and R weighting and is capable of handling some nonminimum phase systems.

6.9 References

Astrom, K.J. (1984). Interaction between excitation and unmodeled dynamics in adaptive control. *Proc. 23rd IEEE CDC*, 1276-1281.

Clarke, D.W. (1984). Self-tuning control of nonminimum-phase systems. *Automatica*, 20, 501-517.

Clarke, D.W. and P.J. Gawthrop (1975). Self-tuning controllers. *Proc. IEE*, 122, 929-934.

Clarke, D.W. and P.J. Gawthrop (1979). Self-tuning control. *Proc. IEE*, 126, 633-640.

Cluett, W.R., S.L. Shah and D.G. Fisher (1986). Robust design of adaptive control systems using conic sector theory. *2nd IFAC Workshop Adapt. Syst.*, Lund, Sweden (to appear in Automatica).

Egardt, B. (1979). *Stability of Adaptive Controllers*, Springer-Verlag, Berlin.

Gawthrop, P.J. and K.W. Lim (1982). Robustness of self-tuning controllers. *IEE. Proc.*, 129, 21-29.

Goodwin, G.C., D.J. Hill, D.Q. Mayne and R.H. Middleton (1985a). Adaptive robust control. University of Newcastle, NSW, Dept. of Elect. and Comp. Eng., Report EE8544.

Goodwin, G.C., D.J. Hill and M. Palaniswami (1984). A perspective on convergence of adaptive control algorithms. *Automatica*, 20, 519-531.

Goodwin, G.C., D.J. Hill and M. Palaniswami (1985b). Towards an adaptive robust controller. *IFAC Identification and System Parameter Estimation*, York, UK, 997-1002.

Goodwin, G.C., P.J. Ramadge and P.E. Caines (1980). Discrete-time multivariable adaptive control. *IEEE Trans. Aut. Contr.*, AC-25, 449-456.

Goodwin, G.C. and K.S. Sin (1984). *Adaptive Filtering Prediction and Control*, Prentice-Hall, Englewood, NJ.

Ioannou, P.A. and K. Tsakalis (1985). A robust discrete-time adaptive controller. University of Southern California, EE-Systems, Report 85-10-1.

Johnson, C.R., B.D.O. Anderson and R.R. Bitmead (1984). A robust, locally optimal model reference adaptive controller. *Proc. 23rd IEEE CDC*, 993-997.

Johnstone, R.M., S.L. Shah and D.G. Fisher (1980). An extension of hyperstable adaptive control to non-minimum phase systems. *Int. J. Contr.*, 31, 539-545.

Kosut, R.L. and B. Friedlander (1985). Robust adaptive control: conditions for global stability. *IEEE Trans. Aut. Contr.*, AC-30, 610-624.

Kosut, R.L. and C.R. Johnson (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Kreisselmeier, G. and B.D.O. Anderson (1986). Robust model reference adaptive control. *IEEE Trans. Aut. Contr.*, AC-31, 127-133.

Lim, K.W. (1982). Robustness of self-tuning controllers. Dept. of Eng. Sci., Oxford University, OUEL Report No. 1422/82.

Martin-Sanchez, J.M. (1974). Contribution to model reference adaptive systems from hyperstability theory. Doctoral dissertation, Universidad Politecnica de Barcelona.

Martin-Sanchez, J.M. (1976a). A new solution to adaptive control. *Proc. IEEE*, 64, 1209-1218.

Martin-Sanchez, J.M. (1976b). Adaptive predictive control system. USA patent no. 4, 197, 576.

Martin-Sanchez, J.M. (1982). A globally stable APCS in the presence of bounded noises and disturbances. *Proc. 21st IEEE CDC*, 761-769; also in *IEEE Trans. Aut. Contr.*, AC-29, 461-464, 1984.

Martin-Sanchez, J.M. (1985). Adaptive control for time-variant processes. *1985 American Control Conference*, Boston, MA, 1260-1268.

Martin-Sanchez, J.M., S.L. Shah and D.G. Fisher (1984). A stable adaptive predictive control system. *Int. J. Contr.*, 39, 215-234.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Praly, L. (1983). Robustness of model reference adaptive control. *Proc. 3rd Yale Workshop*, New Haven, CT, 224-226.

Praly, L. (1984). Robust model reference adaptive controllers, part 1: stability analysis. *Proc. 23rd IEEE CDC*, 1009-1014.

Rohrs, C. (1982). Adaptive control in the presence of unmodeled dynamics. Dept. of Elec. Eng. Comput. Sci., MIT, LIDS Report No. TH-1254.

Rohrs, C.E., M. Athans, L. Valavani and G. Stein (1984). Some design guidelines for discrete-time adaptive controllers. *Automatica*, 20, 653-660.

Sripada, N.R. and D.G. Fisher (1986). An improved least squares estimator. University of Alberta, Dept. of Ch.E., Report 860520.

7. Conclusions and Recommendations

7.1 Conclusions

The contributions of this thesis include a clarification and rationalization of work published by other authors plus significant extensions of these results. The most important result of this work is a formal proof of stability of an adaptive predictive control system in the presence of model-process mismatch (see item #10). However, the following conclusions are presented in a logical order to indicate how the various contributions of this work build upon one another.

1. The global stability results of Kosut and Johnson (1984) for continuous-time adaptive systems in the presence of unmodeled dynamics lead to a strictly positive real (SPR) condition on an operator which is a function of the unmodeled dynamics. Their results are limited by this SPR condition which is almost surely violated due to typically unmodeled, high frequency dynamics. This conclusion justifies considering local rather than global stability results when unmodeled dynamics are known to be present. In this context 'local' means that restrictions are placed on the nature of the external signals and on the initial adaptive parameter error in order to guarantee stability.

2. Analagous results of Ortega *et al* (1985) for

discrete-time adaptive control systems in the presence of unmodeled dynamics lead to a conic sector condition on an operator which corresponds to the SPR condition imposed on its continuous-time counterpart. It is unclear from their results whether or not the conic condition in the discrete-time case is as restrictive as the SPR condition for the continuous-time case when unmodeled dynamics are present.

3. A design approach is presented for discrete adaptive control systems with unmodeled dynamics, based on conic sector conditions, which provides a quantitative measure of the effect of design alternatives such as adaptive gain, model order and sampling rate on stability. These results indicate that the destabilizing effect of high frequency, unmodeled dynamics present in the continuous-time case are lessened by sampling to the extent that global stability results need not be abandoned in the discrete-time case.

4. The nonminimum phase zeros which may arise due to fast sampling of continuous-time plants using discrete adaptive control can cause stability problems even when no unmodeled dynamics are present. Therefore the instabilities that Rohrs (1982) observed should be attributed to a combination of the nonminimum phase zero and the unmodeled dynamics rather than to unmodeled dynamics alone. Other results presented here show that when the sampling interval is chosen to avoid the

nonminimum phase zeros, the system can be either stable or unstable depending on the choice of sampling period, reference input and other design parameters.

5. A modified form of the conic sector stability theorem is developed for discrete adaptive control systems with unmodeled dynamics which accomodates the nonzero initial states of the parameter adaptation algorithms. The proof of this theorem is more straightforward than that originally presented by Safanov (1980).

6. The motivation behind the addition of normalized signals to the parameter adaptation algorithms made by Ortega *et al* (1985) was to avoid an *a priori* signal boundedness assumption in their proof of stability. It is demonstrated here that the particular normalization factor selected by Ortega *et al* (1985) allows for the derivation of a straightforward sector condition on the operator which represents the unmodeled dynamics.

7. The L_2 conic sector stability results of Ortega *et al* (1985) are extended to the L_∞ case (i.e. bounded disturbances and reference inputs). It is demonstrated that out of three parameter adaptation algorithms examined only a least squares algorithm with a forgetting factor less than unity satisfies the conditions required in this extension to the L_∞ result.

8. The predictive control law used by Ortega *et al* (1985) has been augmented to include weighting polynomials on the plant input and output. These weighting polynomials may now be included in Ortega *et al*'s (1985) conic sector formulation.

9. Stability results from previous literature dealing with adaptive predictive control systems (APCS) have been reduced to one concise condition in a new global stability theorem. This condition is not problem specific and depends only on the convergence properties of the *a posteriori* estimation error and the estimated parameters. It is demonstrated that this new theorem is applicable to a broad class of linear, time-varying, stochastic systems.

10. Stability of APCS in the presence of process-model order mismatch has been proven using a normalized parameter estimation scheme, which permits a formal proof that the modeling errors can be treated as a bounded disturbance, and a parameter adaptation stopping criterion. These theoretical results confirm the intuitive expectation that the degree of process-model order mismatch that can be introduced into a stable adaptive control scheme is limited by inherent properties of the process.

11. The assumption of a tuned set of controller parameters and/or the strictly positive real or conicity condition used

by other authors is not required in this approach. However the results presented here show that parameter convergence to such a tuned set is a natural result of the approach used herein.

12. These results have been extended to a self-tuning control (STC) type of adaptive system using a normalized least squares parameter estimation scheme and a stopping criterion. This extended controller is capable of handling some nonminimum phase systems.

7.2 Recommendations

1. Evaluate the effect of other design alternatives such as P and Q weighting on robustness using the design approach presented in Chapter 4.

2. Investigate and compare other normalized estimation schemes in terms of stability and performance.

3. Examine modifications to the least squares algorithm presented in Chapter 6 which retain the necessary convergence properties but prevent the gain of the parameter estimator from going to zero (e.g. constant trace).

4. Examine more closely the parameter a_1 in Chapter 6 and attempt to relate it to other, better understood process parameters.

5. Determine what types of unmodeled nonlinearities can be included in the normalized perturbation signal without violating the boundedness of this signal.

6. Experimentally evaluate the new adaptive control algorithm presented in Chapter 6. This work is currently proceeding at the University of Alberta.

7.3 References

Kosut, R.L. and C.R. Johnson, Jr. (1984). An input-output view of robustness in adaptive control. *Automatica*, 20, 569-581.

Ortega, R., L. Praly and I.D. Landau (1985). Robustness of discrete-time direct adaptive controllers. *IEEE Trans. Aut. Contr.*, AC-30, 1179-1187.

Rohrs, C. (1982). Adaptive control in the presence of unmodeled dynamics. Dept. of Elec. Eng. Comput. Sci., MIT, LIDS Report No. TH-1254.

Safanov, M.G. (1980). *Stability and Robustness of Multivariable Feedback Systems*. MIT Press, Cambridge, MA.

Appendix A: Mathematical Preliminaries

This appendix provides a concise description of the mathematical notation used throughout this thesis. The main concern is with discrete signals which are infinite sequences of real numbers. Each signal may be considered a vector of infinite dimension and represents an element of a set known as a linear vector space.

Norms: Norms may be thought of as a measure of the size of a vector. Let E be the linear vector space. The zero vector in E is denoted by $\emptyset$. The function $\rho:E\rightarrow R_+$ (the set of positive real numbers) is a norm on E if and only if

(i) $x\epsilon E$ and $x\neq\emptyset$ implies $\rho(x)>0$

(ii) $\rho(ax)=|a|\rho(x)$ $\forall a\epsilon R$, $\forall x\epsilon E$

(iii) $\rho(x+y)\leq\rho(x)+\rho(y)$ $\forall x,y\epsilon E$

Given the linear space E and a norm ρ on E, the pair (E,ρ) is called a normed vector space.

L_2-norm: Let $x=(x_1,x_2,\cdots)$. The L_2-norm of x is defined as

$$||x||_2 = \{\sum_{k=1}^{\infty} x_k^2\}^{1/2}$$

L_∞-norm: Let $x=(x_1,x_2,\cdots)$. The L_∞-norm of x is defined as

$$||x||_\infty = \sup|x_k| \quad k\geq 1$$

If these norms exist, the corresponding normed vector spaces are called L_2 and L_∞, respectively. The extension of a space L, denoted by L_e, is the space consisting of those elements x whose truncations lie in L, e.g. x belongs to the extended space L_{2e} if

$$||x||_{2,T} = \{\sum_{k=1}^{T} x_k^{\,2}\}^{1/2} < \infty$$

$\forall T \epsilon Z_+$ (the set of positive integers)

Operator: Operator and relation are used synonymously in this thesis to define a mapping of normed vector spaces.

Gain: Given an operator $H: x \rightarrow y$ where $x,y \epsilon L_e$, suppose that there exist real numbers γ_1 and γ_2 such that

$$||y||_T = ||Hx||_T \leq \gamma_1 ||x||_T + \gamma_2 \quad \forall x \epsilon L_e, \ \forall T \epsilon Z_+$$

The gain of H is the smallest value for γ_1 such that the above inequality holds for some γ_2.

Passivity: Define the scalar inner product $<\cdot|\cdot>$ of two infinite sequences x and y as

$$<x|y> = \sum_{k=1}^{\infty} x_k y_k$$

An operator $H: x \rightarrow y$ where $x,y \epsilon L_{2e}$ is passive if and only if there exists some constant β such that

$$<y|x>_T = <Hx|x>_T \geq \beta$$

Conic Sector: An operator $H: x \rightarrow y$ where $x,y \epsilon L_{2e}$ is:

(i) inside the cone(C,R) if

$$<y-(C-R)x|y-(C+R)x>_T \leq 0 \quad \forall T \epsilon Z_+$$

(ii) outside the cone(C,R) if

$$<y-(C-R)x|y-(C+R)x>_T \geq 0 \quad \forall T \epsilon Z_+$$

(iii) strictly inside the cone(C,R) if for some $\epsilon>0$

$$<y-(C-R)x|y-(C+R)x>_T \le -\epsilon||(x,y)||_T^2 \quad \forall T\epsilon Z_+$$

(iv) strictly outside the cone(C,R) if for some $\epsilon>0$

$$<y-(C-R)x|y-(C+R)x>_T \ge \epsilon||(x,y)||_T^2 \quad \forall T\epsilon Z_+$$

where $||(x,y)||_T^2=(||x||_T^2+||y||_T^2)$

Dissipativeness: An operator $H:x\rightarrow y$ where $x,y\epsilon L_{2e}$ is weakly (Q,S,R) dissipative if and only if there exists a constant β such that

$$<y|Qy>_T + 2<y|Su>_T + <u|Ru>_T + \beta \ge 0 \quad \forall T\epsilon Z_+$$

With $\beta=0$, H is called dissipative.